Degree project

UNIVERSITY OF SKÖVDE

DETECTION OF ROOT BORNE PATHOGEN CAUSING PEA ROOT ROT BY USING MINION SEQUENCING

Master Degree Project in Bioscience

Second Cycle 30 credits

Spring term 2023

Student: Aderemi Olubode

Supervisor: Maria Algerin

Examiner: Magnus Fagerlind

Abstract

Pea (*Pisum sativum L.*) is the most cultivated pulse in the temperate zone, economically important with high nutritional value (high protein content of 20-30%) and relatively low cost. Root rot is its predominant disease, it is referred to as root rot complex because it involves many pathogens. The most important pathogens of pea root rot are fungi and oomycetes. The aim of this project is to assess the possibility of using root samples with known disease severity index (DSI) to identify pea root rot causing pathogens using a MinION device from Oxford Nanopore technologies. DNA barcoding is the use of a standardized segment or region of the DNA to identify the organism to species level by comparison with a reference library. The ITS region is the barcode sequence marker of fungi. In this study, extracted DNA from pea plant roots of six naturally infected pea fields were sequenced. Two different primer pairs; ITS1Catta & ITS4ngsUni (fungi targeted) and ITS1OO & ITS4ngs (oomycota targeted) were used for the polymerase chain reaction (PCR) amplification. Taxonomic identification was done using Kraken2 bioinformatics tool and UNITE reference database. Organisms from fungal phyla Ascomycota, Basidiomycota and Oomycota were recovered from all the samples, but none of the pea rot root specific pathogens (*Sclerotinia sclerotiorum, Pythium ultimum, Thielaviopsis basicola, Fusarium solani, Fusarium oxysporum, Ascochyta pinodella, Aphanomyces euteiches* and *Rhizoctonia solani)* were identified. This project could not validate the use of nanopore sequencing using a MinIon device for the identification of specific pathogens causing pea root rots.

Popular scientific summary

Root is the hidden half of plants and of huge importance in their growth. It also provides a unique environment for numerous microorganisms including bacteria, fungi and archaea because it is rich in a variety of carbon compounds. Fungi are eukaryotes and include microorganisms such as yeasts, moulds and mushrooms. The presence of chitin in fungi cell walls is the main characteristics that differentiates them from plant kingdom. They are the main decomposers in the ecosystems with functions across different fields; medicine, nutrition, biotechnology, veterinary medicine etc. Pea is the small and spherical seed or the pod containing seeds of *Pisum sativum* plant. It is an annual plant cultivated for livestock and human consumption. Root rot is the predominant disease affecting peas and a huge limitation to pea cultivation despite resistance breeding. Disease Severity Index (DSI) is used to denote the severity of diseases. It is defined as the percent area of a sampling unit with symptoms of the disease. The most important pathogens of pea root rot are fungi and oomycetes. These include *Sclerotinia sclerotiorum, Pythium ultimum, Thielaviopsis basicola, Fusarium solani, Fusarium oxysporum, Ascochyta pinodella, Aphanomyces euteiches* and *Rhizoctonia solani.* The commonly selected genetic marker for fungi molecular identification is the ITS region. Internal transcribed spacer (ITS) is the region of non-coding DNA between genes. It is situated between the small subunit ribosomal RNA (rRNA) and large-subunit rRNA genes. The average length in the fungal kingdom is 550 base pairs. Polymerase chain reaction (PCR) is a rapid technique used for making multiple (million or billion) copies of a specific DNA sample. Sequencing using MinION device (Oxford Nanopore Technologies) can be used for fast and accurate identification of pathogenic organisms from plant tissues. MinION is pocket size and potentially offers low cost, high mobility and rapid sample processing with real time result display. This project was to assess the possibility of using pea root samples with known disease severity index (DSI) to identify pathogens causing pea root rot using MinION. DNA was extracted from the root samples of infected pea plant collected from pea fields in Sweden. PCR was used to amplify the ITS region of the DNA and subsequently sequenced with the MinIon. A total of nine samples were sequenced; six PCR amplicons and three DNA samples without amplification. These were selected based on the quality of the DNA sample (purity and quantity). The bioinformatics analysis of the sequenced data was done with the use of Kraken2 taxonomic classification system and UNITE database. Sankey diagrams, providing a display of taxonomic classification to each sample, were generated with the use of Pavian tool. The result showed the presence of expected fungal phyla and other eukaryotes (plants and metazoan). However, none of the pea rot root specific pathogens (*Sclerotinia sclerotiorum, Pythium ultimum, Thielaviopsis basicola, Fusarium solani, Fusarium oxysporum, Ascochyta pinodella, Aphanomyces euteiches* and *Rhizoctonia solani)* were identified.

Abbreviations

DNA - Deoxyribonucleic acid

DSI - Disease severity index

ITS - Internal transcribed spacer

MT - Melting temperature

MW - Molecular weight

NCBI - National Centre for Biotechnology Information

NGS - Next generation sequencing

ONT - Oxford nanopore technology

PCR - Polymerase chain reaction

qPCR - quantitative PCR

RNA - Ribonucleic acid

WIMP - What's In My Pot

Table of content

Introduction

Nutritious and stable food supply is important in the fight against global pandemics. Despite efforts in plant breeding; production of crops with improved resistance to individual pathogens, food security is still threatened by plant diseases caused by multiple pathogens interactions such as root rot in pea (*Pisum sativum L.*) (Lukas et al., 2021). Pea is the small and spherical seed or the pod containing seeds of *Pisum sativum* plant. It can be green or yellow and pea pods are referred to as fruit botanically. The average weight of pea is between 0.1 and 0.36 gram (Duke, 1981). It is an edible seed from the Fabaceae such as; pigeon pea (*Cajanus cajan*), cowpea (*Vigna unguiculata*), and seeds from Lathyrus species. It is an annual plant cultivated for livestock and human consumption. Pea is economically important with high nutritional value (high protein content of 20-30%) and relatively low cost (Peng et al., 2016; Xiong et al., 2018). Peas are susceptible to many diseases with varying severity leading to huge losses. The annual losses vary yearly based on the prevailing local weather conditions. The viral, fungal and bacterial diseases of peas are often disseminated by infected seed, insects, erosion or surface water run-off, manure, farm animals, farm implements, and wind (USDA, 1962). Some of the common diseases of pea are; Ascochyta blight, Bacterial blight, Fusarium wilt, Near-wilt, Root rots, Root knot, Septoria blight, Powdery mildew, Anthracnose, Downy mildew, Pea virus diseases etc (USDA, 1962). The symptoms can be seen on the different parts of the plant; root, stem, leaves etc.

Root is the hidden half of plants and of huge importance in their growth. It absorbs water and nutrients from the soil, serves as nutrient storage and anchors the plant (Lynch, 2007). It is rich in a variety of carbon compounds and as such provide unique environment for numerous microorganisms including bacteria, fungi and archaea. This dynamic community of microorganisms with plant root is referred to as root microbiome. They consist of both beneficial and pathogenic microorganisms which are in constant competition. The pathogenic microorganisms break through the plant defence mechanisms to cause plant diseases (Mendes et al., 2013). The high species diversity of these microorganisms makes it difficult to retrieve them from cultures and it limited the understanding of their assembly until recent advancements with sequencing technologies (Buée et al., 2009).

Root rot is the predominant disease affecting peas (Kumari & Katoch, 2020). It remains a huge limitation to pea cultivation despite resistance breeding as resistance breeding is only effective against individual pathogens (Infantino et al., 2006; Rubiales et al., 2015). Root rot is referred to as root rot complex as it involves many pathogens (Xu et al., 2012), the most important pathogens of pea root rot are fungi and oomycetes. These include *Sclerotinia sclerotiorum* (Figure 1)*, Pythium ultimum, Thielaviopsis basicola, Fusarium solani, Fusarium oxysporum, Ascochyta pinodella, Aphanomyces euteiches* and *Rhizoctonia solani* (Bodah, 2016; Gossen et al., 2016). The ability to survive on plant debris and also form resting structures in the soil further enhance the difficulty in control (Li et al., 2014; Bainard et al., 2017). Earliest symptoms of root rot are underground and not discernible with the plant survival and yields already compromised by the time the symptoms become apparent (Eliane, 2017). Culture based molecular identification of pathogens is limited by species diversity of these pathogens, environmental materials contamination, slow pace and cost. Disease severity index (DSI) is used to denote the severity of disease. Disease severity is defined as the percent area of a sampling unit with symptoms of the disease (Nutter et al., 1991).

Fungi and oomycetes are predominant in pea root rot. Fungi are eukaryotes and include microorganisms such as yeasts, moulds and mushrooms. It encompasses an enormous diversity of organisms with different ecologies, life cycle and morphologies. The diversity of the fungus kingdom is estimated at 2.2 million to 3.8 million species (Hawksworth & Lücking, 2017). The presence of chitin in fungi cell walls is the main characteristics that differentiate them from plant kingdom. They do not photosynthesize, they acquire their food by absorbing dissolved molecules through digestive enzymes secreted into their environment. Fungi are the main decomposers in

the ecosystems; essential role in decomposition of organic matter, nutrient cycling and exchange in the environment. Their function cut across different fields; medicine, human nutrition, biotechnology, veterinary medicine etc. There are nine defined phyla of fungi; Opisthosporidia, Chytridiomycota, Neocallimastigomycota, Blastocladiomycota, Zoopagomycota, Mucoromycota, Glomeromycota, Basidiomycota and Ascomycota (Naranjo et al., 2019). Fungi are responsible for an increasing number of virulent infectious diseases especially in plants; a major global threat to food security. Similarly, oomycetes species are known to be the most destructive pathogen in plants (Hyde et al., 2019).

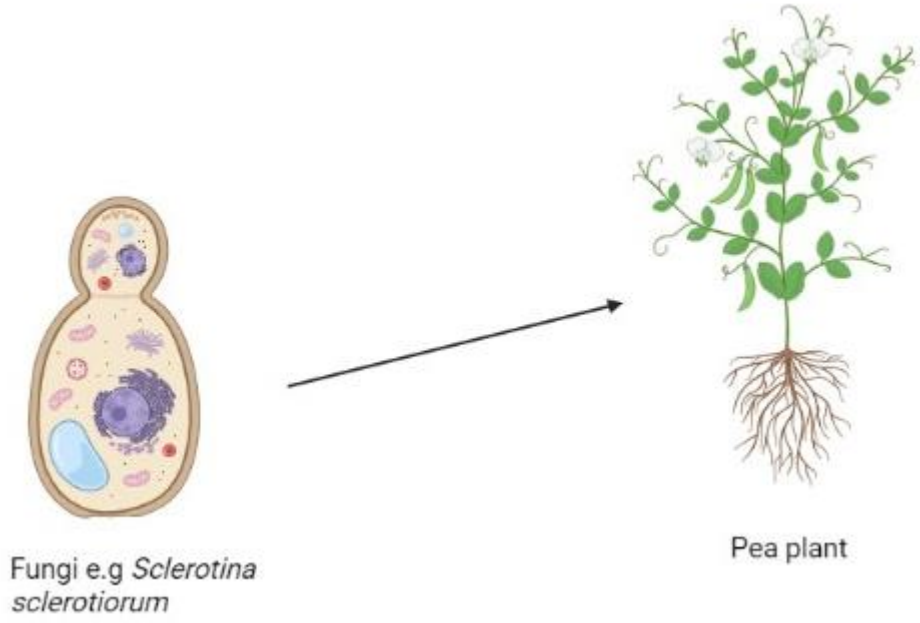

Figure 1. Schematic drawing of a typical pea root rot causing fungi and the Pea plant.

There had had been different approaches to pea root rot disease investigation and control based on technological advancements. Plant disease entails the interaction of plant, pathogen and the environment, this forms the basis of such investigations and control. Greenhouse bio-assay method to predict incidence of root rot based on estimation of potentials in field soil was done by Fink and Buchholtz (1954). This is based on the potential of inoculum on soil samples taken from fields. Pea seeds were planted in the different soil samples in a 6-inch pots within a green house. Percentage of seedlings infected with *A. euteiches* within 30 days was ascertained and subsequently recorded on a scale of 0-100, this is referred to as the disease severity index (Johnson, 1953). Identification of fungi often involve investigation in the laboratory after they have been harvested from the field, this makes accurate and precise identification challenging. This is in comparison with other multicellular eukaryotes as some of the features are only discernible in the fields (Beakes & Thines, 2016).

Further technological advancement had birthed different approaches to fungal studies. There is the use of the whole genome in fungi studies (Gladieux et al., 2015), Wibberg et al. (2020) implemented the genome-based approach described by Qin et al. (2014) to resolve species complexes in eukaryotes using percentage of conserved proteins (POCP) analysis. Also, there is the molecular approaches where gene regions are used for detection of microorganisms; 16S rRNA for prokaryotes and 18 rRNA for eukaryotes (Paulay, 2005). The ITS region is also used as a genetic marker (Wurzbacher et al., 2019). It is a commonly selected genetic marker for fungi molecular identification (Nilsson et al., 2015). The ITS region has the advantage of ease of amplification, primer annealing, wide spread use and large barcode gap (Begerow et al., 2010). There are several designed primers for the ITS regions, these primers are for direct amplification of fungal DNA from samples with multiple DNA sources such as plant tissues and soil (Gardes & Bruns, 1993). DNA barcoding is the use of a standardized segment or region of the DNA to identify

the organism to species level by comparison with a reference library. Amplification prior sequencing is necessary because root is a complex substrate; contains multiple sources of DNA. Amplification of the DNA sample is done using polymerase chain reaction (PCR).

PCR is a laboratory technique used for the rapid production (or amplification) of several (millions to billions) copies of a specific DNA segment to enhance detailed subsequent studies. It entails the use of short synthetic DNA fragments that are referred to as primers, these select a segment of the genome to be amplified. Afterwards, there are multiple rounds of DNA synthesis for the segment to be amplified. Fungi specific primers directly amplify the fungal DNA alone which helps to separate it from other DNAs (Gardes & Bruns, 1993). The primers used in this project were used to amplify the ITS regions of fungi and oomycetes species; forward primer ITS1catta and reverse primer ITS4ngsUni for fungal ITS, ITS1OO and ITS4ngs primers as forward and reverse primers respectively for oomycetes (Tedersoo et al., 2014).

Sequencing using MinION device (Oxford Nanopore Technologies) can be used for fast and accurate identification of multiple pathogenic organisms from plant tissues (Loit et al., 2019). DNA sequencing is the use of technology to determine the order of arrangement of the four nucleotide bases. It has undergone several advancements since its initial use or adoption by Frederick Sanger. Primers extension strategy was used to develop what is known as first generation sequencing (Sanger et al., 1977). Afterwards was the improvement based on sequencing speed and lower cost, referred to as second generation sequencing; SOLiD, 454, Ion Torrent and Illumina (Kchouk et al., 2017). Next generation sequencing (NGS) is broadly the DNA sequencing technology that is advanced more than Sanger sequencing. Also, real time sequencing data is generated in NGS as each nucleotide is added to the nascent strand of DNA (Datto & Lundbla, 2016). The design of third generation sequencing as seen in Oxford Nanopore technologies (MinION, GridION and PromethION) and Pacific Biosciences (RSII and Sequel) was to further reduce cost, increase speed and bring possibility of large amount of sequences from many samples without compromise on accuracy of the reads. It has improved read length of about 20,000-1,000,000 bp compared to approximately 800 bp and 400 bp of Sangers and NGS respectively (Jain et al., 2016). Nanopore sequencing is a third-generation sequencing approach for DNA or RNA (Niedringhaus et al., 2011). One of the available devices to perform nanopore sequencing is MinION (Oxford Nanopore technologies). It is a portable 90 g device that can plug into any computer with a standard USB 3.0 port and offers different applications as seen in whole genome sequencing (Jain et al., 2016), genotype analysis of cancer (Euskirchen et al., 2017), etc. Oxford nanopore offers platforms for further analysis; EPI2ME and WIMP, these provide analysis of the generated FASTQ files. EPI2ME is an analysis platform for both cloud-based and local data; it provides access to data analysis workflows while WIMP (What's In My Pot) is a workflow that utilises nanopore sequencing reads from metagenomic samples for species identification and quantification (Oxford nanopore technologies). Similarly, Kraken2 can be used for taxonomic classification system for high accuracy and speed (Wood et al., 2014). Also useful is the UNITE database (curated sequence database) with about 2.5 million of curated fungal ITS sequences which corresponds to over 100,000 species with 98.5% identity (Schoch et al., 2012).

Aims and objectives

The aim of this project is to assess the possibility of using root samples with known disease severity index (DSI) to identify pea root rot causing pathogens using MinION from Oxford Nanopore technologies. The objectives are: extract genomic DNA from naturally infected pea root samples with varying and known disease severity index. PCR amplification of the ITS regions of the extracted genomic DNA with the use of appropriate primers. Carry out Nanopore sequencing using MinIon from Oxford Nanopore Technologies, analysis of the generated data and make a comparison between the pathogens detected in the different root samples to the different disease severity index (DSI). Apart from the direct benefit to farmers with reduction in losses to root rot

diseases through prompt diseases diagnosis, it will also form a basis that can be replicated in the diagnosis of similar fungal diseases in both plants and animals.

Materials and method

Biological materials

The samples used for this project were root samples of infected pea plants that were collected from five pea fields in Sweden. These were stored at -20 °C in the laboratory prior DNA extraction from them. The samples have different known DSI that was analysed and calculated using fundus (Table 1).

Table 1. The root samples with DSI.

Sample ID	DSI
15020	66
ORSMARK2	36
14735	28
FRANK K	25
15015	19

DNA Extraction

DNA was extracted from the root samples of infected pea plant using E.Z.N.A. ® SP Plant DNA Kit (OMEGA Biotek). The kit manufacturer's instruction was followed for the DNA extraction. The root samples were grinded in liquid nitrogen for tissue disruption as recommended in the protocol. The DNA extraction was carried out using 100 mg each of the root samples and 50 µl of elution buffer was used to elute the DNA.

DNA quantification

Qubit 4.0 (Invitrogen) equipment with dsDNA HS assay kit (Thermofisher Scientific) was used for quantification (concentration) of the extracted DNA. Sample purity (determine the absorbance ratios A260/280 and A260/230) was ascertained using a DeNovix DS-11 spectrophotometer (Thermofisher Scientific).

PCR and electrophoresis

All the five samples were used during the optimization PCR. It was done in a 20-µL PCR mixture (Table 2) with five annealing temperatures (55 °C, 58.5 °C, 60 °C, 62 °C and 63.2 °C). The samples were diluted to a concentration of 15ng/µl. Phusion High Fidelity DNA polymerase (New England Biolabs) and Phusion High Fidelity PCR kit were used for the polymerase and programme respectively. The primers; forward primer ITS1catta and reverse primer ITS4ngsUni were used to target the fungal ITS region. Similarly, forward primer ITS1OO and reverse primer ITS4ngs were used for oomycetes ITS region (Appendix 1). There were several repeats of the optimization PCR because of the non-conclusive result as seen on agarose gel electrophoresis. Several changes were instituted; another set of annealing temperature (57.6 °C, 60.4 °C, 61.4 °C and 66.1 °C), different DNA template amount, that is, both lower and higher than 105 ng (25 ng, 50 ng, 100 ng and 200 ng) were checked. Also, there were changes in the concentrations of the primers (0.5 µM to 1µM) and change in the DNA polymerase from Phusion High Fidelity DNA polymerase (New England Biolabs) to Phusion Hot Start II High-Fidelity DNA polymerase (ThermoScientific). Summarily, six different PCRs were carried out for optimization.

Table 2. PCR optimization, 20-µl reaction

Component	20µl Reaction	Final concentration (µl)
Water	1µl	-
Forward primers	1 µl	0.5 µM
Reverse primers	1 µl	0.5 µM
Template DNA	7 µl	15 ng/µl
Master Mix	10 µl	2X

The PCR amplification of the extracted DNA was carried out using 105 ng of template DNA with the use of a BIO-RAD IPTC-200 thermal cycler. The PCR was run for 35 cycles at initial denaturation temperature of 98 °C for 30 seconds, denaturation at 98 °C for 10 seconds, annealing at 55 °C for 30 seconds, extension at 72 °C for 30 seconds and final extension at 98 °C for 10 minutes.

The amplicons (10µL) with 2µL of 6X purple gel loading dye (New England Biolabs) were run on 0.8% agarose gel at 100 volts for 1 hour. 1X TAE buffer stained with 1X GelRed® (Biotium) according to manufacturer's instruction was used in running the gel electrophoresis and a total of 10 gel electrophoresis from the five samples (with both primer pairs).

Cleaning

This was done while troubleshooting the cause of the inconclusive result on agarose gel electrophoresis. The extracted DNA from sample 15015 was cleaned with the use of 3M Sodium Acetate at pH 5.2 (this was prepared according to the preparation protocol of AAT Bioquest) to enhance the DNA quality (purity). Sample 1505 was selected to test for cleaning because it has the highest concentration among all the samples. The cleaned sample was used for another annealing temperature optimization PCR (using 55 °C, 58.5 °C, 60 °C, 62 °C and 63.2 °C).

MinION sequencing

Library preparation

A total of nine samples selected from three fields (Appendix 2) out of the initial five fields were used for MinIon sequencing. A total of 37 ng of DNA (equivalent to 120 fmol) was used per sample. Online calculator (NEBioCalculator) was used for the conversion of ng to fmol. Also, 31 ng (equivalent to 100 fmol) was used for the native barcode ligation. The three samples used are 15020, ORSMARK2 and 14735. The selection of these three samples was based on the quality (absorbance ratios; A260/280 and A260/230) and quantity available from the extracted DNA. Three samples each were selected from PCR primer pairs ITS1catta & ITS4ngsUni and ITS1Oo & ITS4ngs. The last three samples that made up the total number are the extracted DNA without amplification. The Oxford Nanopore technologies native barcoding protocol was used alongside native barcoding expansion kit 1-12 (EXPNBD104) and ligation sequencing kit (SQK-LSK109) in the library preparation (Oxford Nanopore technologies). The instruction in the protocol was followed without amendment. The SFB (Short fragment buffer) was used for the adapter ligation and clean up step though short fragment buffer (SFB) and long fragment buffer (LFB) were available in the kit. The buffers are meant to retain all sizes of DNA fragments and to enrich for DNA fragments of 3 kb or longer respectively. Only the SFB (Short fragment buffer) was used for the adapter ligation and clean up step since the expected fragment size was 400-900 base pairs (500 bp in length).

Sequencing and basecalling

The selected nine samples (Appendix 2) were sequenced with the FLO-MIN106 (R9.4.1) flow cell at room temperature. Quality checking of the MinION hardware and flow cell status was carried out prior loading. Approximately 800 active pores were available for sequencing. The MinION was connected to a laptop with internet access and sequencing left for 20 hours. It was stopped at 20 hours after it was deduced to have obtained sufficient amount of sequences. The FAST5 to FASTQ file generation was set to automatic and executed with the use of the in-built Guppy basecalling software (ONT). Also, minimum barcoding score was set at 60 and high base calling accuracy was selected for the Guppy command line software. Lastly, the reads were placed into passed and failed bins accordingly based on the default setting (Phred score). The flow cell was washed and storage buffer added after sequencing. Flow cells were stored for later use at 4°C.

Sequence analysis

The Bioinformatics analysis of the data generated from sequencing was executed with the use of Kraken2 taxonomic classification system (Wood et al., 2014) and UNITE database to cater for the probable undescribed or barcoded species in the samples (Tedersoo et al., 2022). The output data from the sequencing which was in the FASTQ format were compressed to generate fastq.gz files with the use of command cat*. Kraken2 software provides a fast taxonomic classification of metagenomic sequence data (Wood et al., 2019). It retrieves the lowest common ancestor based on multiple best hits in metagenomics and metabarcoding (Wood et al., 2019). UNITE database contains eukaryotic ITS database (Abarenkov et al., 2021). Also, it is the largest database as it contains data directly submitted and curated data from International Nucleotide Sequence Databases consortium (INSDc) (Nilsson et al., 2018). Sankey diagrams were generated with the use of Pavian tool (Ondov et al., 2011). The Pavian tool explored the database to provide a display of the taxonomic classification of each sample.

Results

DNA extraction

DNA extracted from the five root samples (that is, the root samples that were obtained from infected pea plants) were quantified (Table 3). The quantification result showed variation in the DNA concentration among the five samples while the purity (A260/280 ratio) was approximately similar. The A260/230 ratio was within the range of 0.80 to 1.30. Sample 15015 had the highest concentration of 53.80 ng/µl while sample FRANK K had the lowest concentration of 15.90 ng/µl. All the samples had purity value (A260/280 ratio) that is close to 1.8. The yield of the DNA with the use of E.Z.N.A. ® SP Plant DNA Kit (OMEGA Biotek) was good considering the subsequent use, which is for PCR amplification and MinIon sequencing.

Table 3. Quantification results for extracted DNA from the pea root samples

Sample ID	DSI	Qubit Concentration (ng/µl)	Absorbance 260/280	Absorbance 260/230
15020	66	17.90	1.70	0.80
ORSMARK2	36	38.20	1.85	1.30
14735	28	17.40	1.87	1.00
FRANK K	25	15.90	1.85	0.88
15015	19	53.80	2.01	0.61

PCR and electrophoresis

PCR optimization was done to ascertain the most suitable annealing temperature using the DNA from the five samples. Five annealing temperatures (55 °C, 58.5 °C, 60 °C, 62 °C and 63.2 °C) were tested using both primer pairs ITS1catta & ITS4ngsUni and ITS1OO & ITS4ngs. The outcome of the PCR was subsequently run on agarose gel electrophoresis to ascertain the optimal PCR conditions. There were no amplification as distinct bands were not seen on the agarose gel (Not shown). As such, no decision could be made on the optimal annealing temperature. Thorough troubleshooting of the PCR and agarose gel electrophoresis was done. This necessitated several changes and repetitions of optimization PCR. These changes include; PCR optimization for annealing temperature was repeated with another set of temperatures (57.6 °C, 60.4 °C, 61.4 °C and 66.1 °C) using the five samples with the primer pairs. Also, subsequent optimization PCRs with variations in DNA template amount (25ng, 50ng, 100ng and 200ng), different concentrations of the primers (0.5 µM, 1µM), and change of DNA polymerase (Phusion High Fidelity DNA polymerase (New England Biolabs) was changed to Phusion Hot Start II High-Fidelity DNA (ThermoScientific). However, the results on agarose gel were not different from the initial result. Similar result was obtained on agarose gel (Figure 2) with the PCR amplification done at annealing temperature of 55 °C, Phusion Hot Start II High-Fidelity DNA (ThermoScientific) polymerase and both primer pairs ITS1catta & ITS4ngsUni (Lane 2-6) and ITS1OO & ITS4ngs (Lane 10-14).

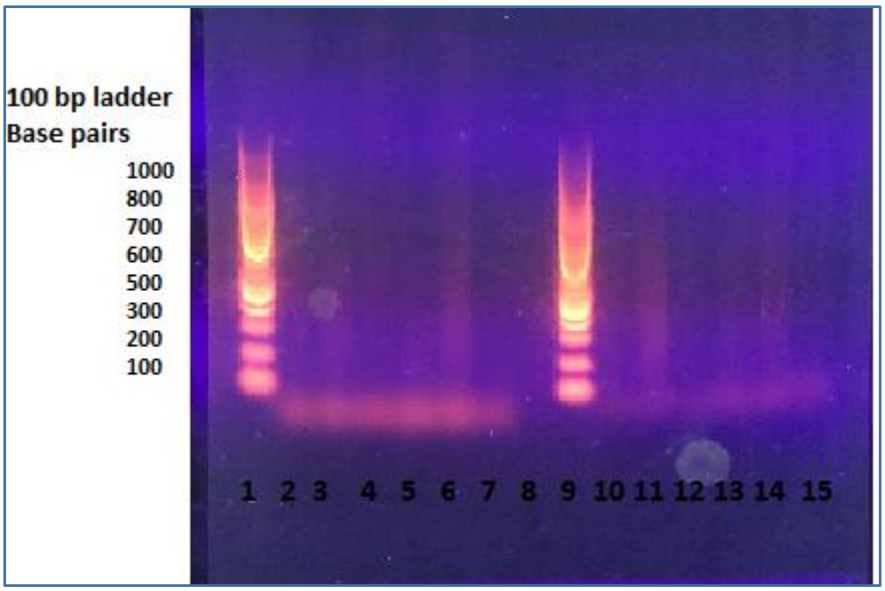

Figure 2. Gel image of the PCR amplification at 55 °C of the five DNA samples with the use of primer pairs ITS1catta & ITS4ngsUni (lane 2-6) and ITS1OO & ITS4ngs (lane 10-14). Lane 1 and 9 are 100bp DNA Quick load ladder (New England Biolabs). Lanes 2,3,4,5 and 6 are for samples 15020, ORSMARK 2, 14735, FRANK K and 15015 respectively (with primer pairs ITS1catta & ITS4ngsUni) while lane 10, 11,12,13,14 and 15 of samples 15020, ORSMARK 2, 14735, FRANK K, and 15015 respectively (with primer pairs ITS1OO & ITS4ngs).

DNA cleaning

The last troubleshooting to improve optimisation PCR was the cleaning of the sample. One of the DNA samples (sample 15015) was cleaned with the use of 3M Sodium Acetate at pH 5.2 (prepared according to the preparation protocol of AAT Bioquest). Quantification was done after the cleaning with the A260/280 and A260/230 ratios before and after cleaning almost the same; (2.01 and 1.81) and (0.61 and 0.61) respectively. Subsequent optimisation PCR carried out alone with this sample (15015) did not give any change in outcome on agarose gel.

MinION sequencing

This entails the two main steps of; library preparation step as well as priming and loading the flow cell step. There were a total of nine samples selected for sequencing (Appendix 2). The nine samples were gotten from categorising the extracted DNA into two categories; amplicons and non-amplified. The amplicons were further divided into two categories based on the two primer pairs (ITS1catta & ITS4ngsUni and ITS1Oo & ITS4ngs). Three samples each were selected from the original five samples from the created categories (two categories from the two primer pairs and the genomic DNA (unamplified) category. These three samples were selected based on most suitable quality (purity) and concentration of the extracted DNA suitable for sequencing since the gel image (Figure 2) was not conclusive to ascertain the qualities of the amplicons. A total amount of 37 ng of DNA, which is equivalent to 120 fmol, was used for the initial sequencing step (end preparation) where a range of 100-200 fmol of DNA was required according to the protocol. The quantification result of the end-prepped DNA done with 1 µl of eluted sample (Appendix 3) showed low concentration. The low concentration resulted in the use of 31 ng (which is equivalent to 100 fmol) for the native barcode ligation step. This is the lower value (100 fmol) of the recommended range (100-200 fmol) that was used. The quantification result after the native barcode ligation step showed a further reduction in concentration of all the samples (Appendix 2). The adapter ligation and clean up step concludes library preparation. The quantification result

obtained for 1 µl of the adapter ligated DNA (Appendix 4) was satisfactory. The pooled barcode DNA samples had 1.44 ng/µl in 15 µl (equivalent to 69.9 fmol) and 7.8 ng (25 fmol) was loaded unto the flow cell, this is within the range of the recommendation (5-50 fmol).

Sequencing performance

The result of the sequencing performance of the nine samples showed the percentage of classified reads across the samples with a range of 1.0-4.9% and 3.2-9.6% for Kraken2 and UNITE respectively (Appendix 5). Highest percentage of unclassified reads (98.7-99.0% and 95.6-96.8% for Kraken2 and UNITE respectively) was seen in the unamplified DNA samples despite having higher number of total reads (Figure 3a-c) compared to the amplified DNA samples which had higher percentage of classified reads. Sankey visualisation output with the use of Pavian tool was achieved for all the nine samples. The result is similar across the samples and the typical illustration as seen across all the samples is shown with sample ORSMARK2 (Figure 3a-c); amplified with primer pair ITS1Catta & ITS4ngsUni (Figure 3a), ORSMARK2 amplified with primer pair ITS1OO & ITS4ngs (Figure 3b) and ORSMARK2 unamplified DNA (Figure 3c). All the samples as typically shown with sample ORSMARK2 (Figure 3a-c) have organisms identified from Kingdom Fungi, Metazoa and Viridiplantae. Kingdom Fungi have the highest number of identified organisms (over 60%) while kingdom Metazoa have the lowest (4%) number of identified organisms.

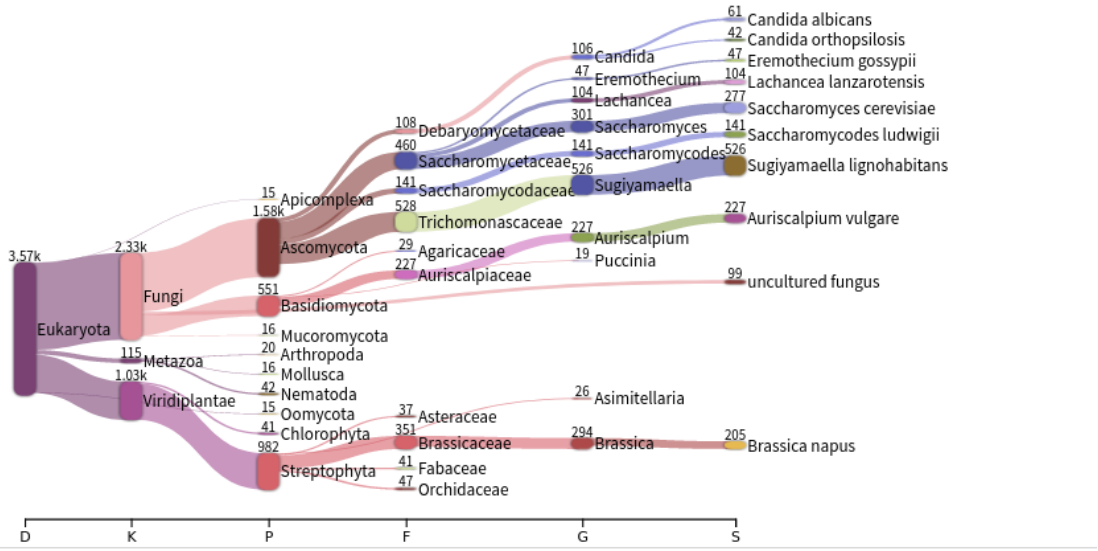

Figure 3a. Taxonomic identification of sample ORSMARK2 amplified with primer pair ITS1Catta & ITS4ngsUni with the use of UNITE database. The numbers on the tree depicts the number of reads for each group. The taxonomic groups are denoted by letters; D-domain, K-kingdom, P-phylum, F-family, G-genus, S-specie.

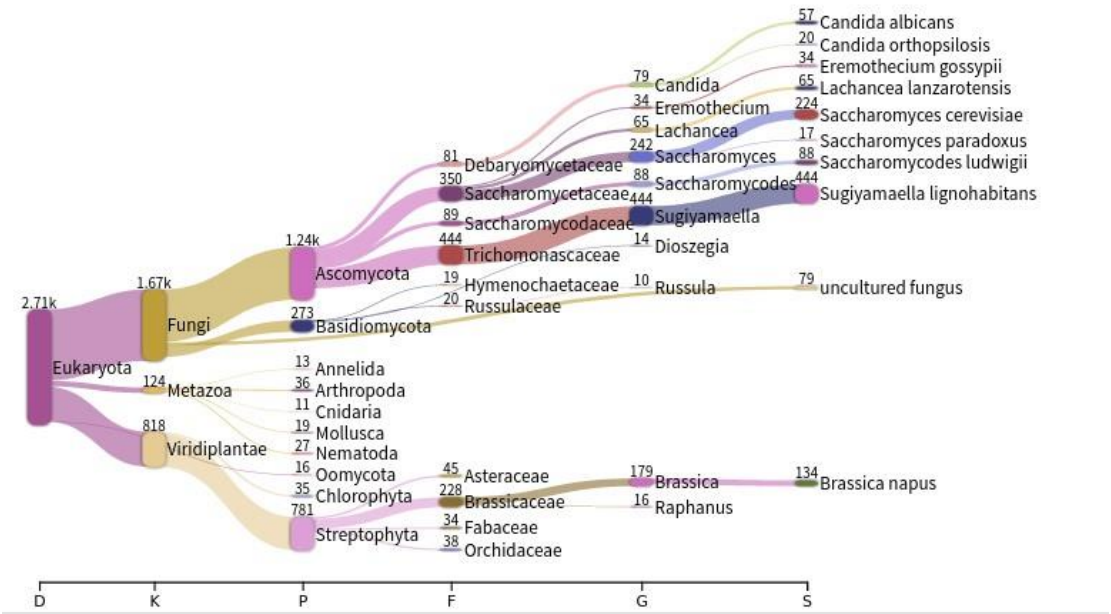

Figure 3b. Taxonomic identification of sample ORSMARK2 amplified with primer pair ITS1OO & ITS4ngs with the use of UNITE database. The numbers on the tree depicts the number of reads for each group. The taxonomic groups are denoted by letters; D-domain, K-kingdom, P-phylum, F-family, G-genus, S-specie.

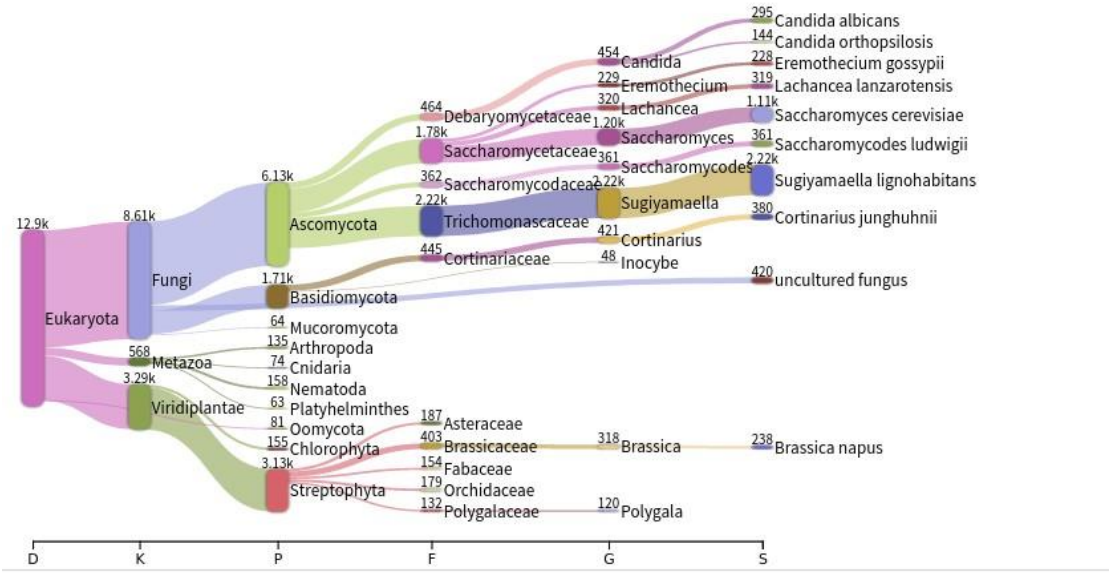

Figure 3c. Taxonomic identification of sample ORSMARK2 (unamplified DNA) with the use of UNITE database. The numbers on the tree depicts the number of reads for each group. The taxonomic groups are denoted by letters; D-domain, K-kingdom, P-phylum, F-family, G-genus, S-specie.

Discussion

The aim of this project was to assess the possibility of using root samples with known disease severity index (DSI) from pea fields in Sweden to identify pea root rot causing pathogens using MinION from Oxford Nanopore technologies. Also, to compare the impact of the different DSI values on the pathogens identified from the root samples.

The DNA extraction was done using E.Z.N.A. ® SP Plant DNA Kit (OMEGA Biotek) according to the manufacturer's instruction. This kit was chosen based on its efficacy, wide application in DNA extraction from plants tissues and recommendation from similar research works (Zhu et al., 2006; Wallenhammar et al., 2012). It also offers time saving advantage as well as direct downstream applications since precipitation using isopropanol or ethanol is eliminated in the procedure and the purified DNA is eluted in water or low ionic strength buffer. E.Z.N.A. ® SP Plant DNA Kit (OMEGA Biotek) system is based on specific and reversible binding of the DNA or RNA at optimized condition in a matrix while proteins and other contaminants are removed. DNA extraction method is an integral part of molecular analysis of plant tissues especially root samples as it contains soil and other organic particles. The objectives of DNA extraction must be met; complete cell lysis and extraction of DNA free from all forms of inhibitors. The result, in terms of the concentration of the extracted DNA in this study is good or acceptable while that of sample purity is not thoroughly acceptable (Table 3). As it is seen from the result (Table 3), A260/280 ratios of all the samples range from 1.80 to 2.01 while the optimal range for purity is from 1.80 to2.00. Lower ratio suggests a possibility of contamination from protein (Wilfinger et al., 1997; Schalamun et al., 2019). This result is similar to the works of Zhu et al. (2006), Rantala-Ylinen et al. (2011) and Wallenhammar et al. (2012). However, the A260/230 ratio, which is used as a secondary measure of nucleic acid purity was within the range of 0.61 to 1.30 (Table 3). This is lower than the expected range of 2.0-2.2 and indicates presence of contaminants such as phenol, which may interfere with downstream applications (Arseneau et al., 2017). The A260/230 ratio assesses possible contamination with substances that absorb strongly at 230 nm. Examples of such substances are; residual phenol, guanidine thiocyanate, or other contaminants commonly present in nucleic acid extraction methods (Sharma et al., 2008). The result of the cleaning with the use of 3M Sodium Acetate at pH 5.2 (prepared according to the preparation protocol of AAT Bioquest) did not give the desired impact; the A260/280 ratio before and after cleaning was almost the same, 2.1 and 1.81 respectively while A260/230 did not change at all from 0.61 which is similar to the works of Bongam (2022). Cleaning after DNA extraction with effective purification method like the use of GE Healthcare kit is imperative for the removal of unwanted substances and contaminants (Smith et al., 2015; Bongam 2022). Furthermore, the concentrations of the extracted DNA from the root samples varied (Table 3). This might be due to the difference in the age of the root samples from the different fields as there was no record or consideration for age of the samples. Plants produce and exude different metabolites through their roots and this can affect the microbes or organisms that cluster around the roots even before reaching the root surface where they will be confronted by the immune system of the plant (Sasse et al., 2017). The age and the stage of development of the plant influence exudation which ultimately affect the microbes proliferating around the roots (Chaparro et al., 2013). Also, according to Handayani et al. (2021), the age of plant sample has direct impact on the concentration of DNA extracted from plant materials. Assessment and taking into cognizance of age of the root or plant sample prior DNA extraction in future studies will help to prevent significant variations in the quantification (concentration) results (Chaparro et al., 2013; Handayani et al.,2021).

The result on agarose gel showed that there was no amplification; the band positions on agarose gel were expected to match the expected number of base pairs, 400-900 bp. Two different primer pairs, ITS1Catta & ITS4ngsUni for the fungal DNA ITS region (Ascomycota, Basidiomycota and Zygomycetes) and ITS1OO & ITS4ngs for oomycetes (Loit et al., 2019) were used in the

amplification PCR. This is contrary to the works of Loit et al. (2019) and Bongam (2022) who amplified fungi using ITS1Catta & ITS4ngsUni at annealing temperature of 55 °C with visible amplicons and more product recovery. Several adjustments were made in line with the works of Lorenz (2012) on annealing temperature and template DNA (amount and quality) as possible limitations to the optimum performance of the PCR. These changes include different amounts of template DNA, annealing temperatures, cleaning of the DNA and the use of already prepared (customized) mastermixes). However, the changes did not affect the result as exactly the same outcome was observed on the agarose gel electrophoresis (there was no distinct band) on several repeats. This might be due to the contamination of the extracted DNA with probable presence of inhibitors; humic and fulvic acid. The quality of extracted DNA from preserved plant and fungal specimens is commonly affected or compromised by the presence of plant-based PCR inhibitors and microflora present as contaminants at the time of collection (Choi et al., 2015). Similar result was observed by Smith et al. (2020) who tested 13 extraction kits including E.Z.N.A. ® SP Plant DNA Kit (OMEGA Biotek) in order to identify a reliable DNA extraction protocol for use in producing high quality sequences for fungal taxonomy. Organic particles or humic substances alongside soil in the root samples contains inhibitors, typical examples of such inhibitors are humic acid and fulvic acid (Tebbe & Vahjen, 1993). Inhibition of DNA enzymatic manipulations by humic acid and fulvic acid occur through interference with the enzymatic activities of DNA polymerase during DNA enzymatic manipulations or reactions when it is present in the extracted DNA (Tsai et al., 1992). It also occurs by binding directly to the DNA (Sutlovic et al., 2008). However, humic acid will not bind to DNA under regular conditions of PCR (Sidstedt et al., 2020). It is suggested that the phenolic structures of humic substances are possibly responsible for inhibition of DNA polymerase because the same level of inhibition is observed in different soils and standardized preparations of humic acid. Also, it has similar characteristics (size and charge) to DNA which affects DNA quantification (purity and concentration) and measurement (Albers et al., 2013). The insufficient time hindered the testing of other cleaning methods as a lot of time was spent in the troubleshooting of the PCR and gel electrophoresis. Also, dilution of the DNA template was not done as it significantly reduces estimate of fungal richness and diversity or the target DNA template number per sample. This can lead to shift in composition of the fungal community (Sarah et al., 2018). The effective method to remove humic acid from DNA was demonstrated by Robe et al. (2003) with the use of Sepharose resins. This can be considered to improve the outcome of future studies.

According to Gardes and Bruns (1993), amplification prior sequencing is essential because root is a complex substrate that contains multiple sources of DNA. The fungi specific primers used to amplify the fungal DNA help to separate it from other DNA that might be present in the sample (Gardes & Bruns, 1993). However, the MinION device from Oxford Nanopore Technologies, like other third generation sequencing technologies, offers the potential to sequence native DNA and RNA without amplification requirement. MinIon sequencing entails different and significant steps; end preparation, native barcode ligation, adapter ligation with clean up as well as priming and loading of the flow cell. The primer pairs ITS1Catta & ITS4ngsUni and ITS1OO & ITS4ngs were meant for amplification of the fungal ITS regions (Loit et al., 2019). ITS region is broadly used in DNA barcoding and metabarcoding analyses of fungi because of its multiple copy numbers, possibility of designing both universal and fungal primers as well as high species level resolution in most of the groups (Nilsson et al., 2018; Schoch et al., 2012). The most important specific pathogens causing pea root rot (*Sclerotinia sclerotiorum, Pythium ultimum, Thielaviopsis basicola, Fusarium solani, Fusarium oxysporum, Ascochyta pinodella, Aphanomyces euteiches* and *Rhizoctonia solani*) are classified into three phyla (Ascomycota, Oomycota and Basidiomycota). Organisms of fungi phyla Ascomycota, Basidiomycota, Oomycota as well as others from Metazoa and Viridiplantae kingdoms were recovered in all the samples. This is similar to result obtained by Loit et al. (2019) who amplified fungi from conifer (Pinaceace) needles, potato (*Solanum tuberosum*) leaves and tubers which resulted in 74.2 % of fungi and 26.1 % of others. Variations were observed in the classified reads among the different samples. Similar studies on fungi and oomycetes ITS which made use of UNITE database resulted in reads constituted mainly of fungi,

metazoan and bacteria (Tedersoo et al., 2014). The results from the nanopore sequencing showed that the unamplified DNA had the highest percentage of classified reads while primer pair ITS1OO & ITS4ngs had the lowest percentage of classified reads (Appendix 5). This confirms the potentials of the third-generation sequencing technologies to sequence native DNA and RNA without amplification requirement. Also, the lower percentage of classified reads compared to that of the unamplified DNA might be a form of bias as different primers can amplify different taxa preferentially (Sipos et al., 2007) The organisms identified in the three categories of the samples are similar; most featured are organisms of the genera Candida, Eremothecium, Lachances, Saccharomyces, Saccharomycodes and Brassica. Brassica is from the kingdom Viridiplantae while others are from the Fungi kingdom (basically yeast). *Brassica napus* is rapeseed. Rape seed is bright-yellow flowering plant with an oil rich seed. It is basically cultivated for its oil which is of immense use. It is often grown together with peas in Europe as an annual break crop in three to four-year rotations with other crops especially cereals such as barley and wheat (Alford, 2008). This might be the reason for its occurrence and identification in this study. Yeasts are found in the soil and decomposing plant materials, they dominate fungal succession during plant decay (Martin et al., 2021). They are chemoorganotrophs; organisms that use organic compounds instead of sunlight as their source of energy. It is a form of nutrition that involves the processing or decomposition of decayed organic matter. Some of the species of yeast metabolize and utilise pentose sugars such as ribose (Barnett, 1975). Yeasts are of immense usage and benefits in industries as seen in the production of ethanol, D-xylitol (Chen et al., 2010) and other non-alcoholic beverages. In nutrition, for the production of food supplements and probiotics (Johnson et al., 2012). They are also used in bio factories and scientific research (Botstein et al., 2011). Some are also pathogenic in plants and animals (Deacon, 2005) e.g Eremothecium, Candida etc. *Eremothecium gossypii* identified in all the samples is one of the pathogens causing stigmatomycosis (Ashby & Nowell, 1926). It occurs in many crops such as cotton, citrus, soybean, pomegranate etc. In cotton, it typically affects the development of hair cells which makes them to dry out and collapse.

According to Alberta Pulse Growers (2014) and Chatterton et al. (2015), pathogens causing root rots are difficult to identify in severely damaged or dead plants because of the presence of other organisms that feed on the decaying tissue as well as pathy distribution of the pathogens. This might suffice for the inability of this study to detect the most important specific pathogens causing root rots in peas as most of the organisms identified are essentially saprophytes. The root samples used for this study were not from fresh plants. Future studies should consider harvesting samples from fresh plants as pointed out in this study as this improve the possibility of detecting the specific pathogens of root rot in peas.

Conclusion

The study cannot sufficiently confirm the possibility of using the primer pairs ITS1Catta & ITS4ngsUni and ITS1OO & ITS4ngs in PCR amplification prior Nanopore sequencing using MinIon device (ONT) for the identification of the pathogens of pea root rots. The sample collection, DNA extraction, cleaning protocols and PCR optimisation will require more investigation to ascertain the potential of this technology in identifying specific pathogens of pea root rot. The result of this study gives supporting information to the significance of the root sample integrity (age and level of decomposition) in the detection of pea root root specific pathogens. This will prevent agricultural losses from pea root rot diseases by enhancing precise and prompt detection of the pathogens.

Ethical aspects and impacts on the society

All genetic materials used in this project were recovered from pea roots collected from five different pea fields in Sweden. As such, no ethical approval was required as the details or personal information of the farmers were not attached to the samples. Also, this project can possibly create a quicker and cost-effective option to the detection of root rot in pea especially as its diagnosis is only beneficial prior symptoms manifestation. Considering the economic impact of pea in food security for man and animal, root rot disease early diagnosis is imperative. The project can similarly form a template for other plant and animal fungal disease diagnosis using nanopore sequencing tool.

The methodology of this study is reproducible and the tools used for data processing are openly available publicly or on purchase of the MinIon device. The University of Skövde made available the funds for this project. Pea is an annual plant cultivated for both livestock and human consumption. Root rot is the predominant disease affecting peas (Kumari & Katoch, 2020) and remains a huge limitation to pea cultivation. Earliest symptoms of root rot are underground and not discernible easily, as such, a fast and accurate method for detection will be of immense benefit. This study adds to the information available on the use of MinIon for pathogen detection studies and research activities; it is a judicious and economical use of funds.

Root rot in peas pose a big problem on food production and quality. It is a complex disease that affect the portion of the developing plants that is below the ground. The causative organisms are borne by the seed or soil and can infect the plant at any stage. Effects of fungal pathogens are also suffered by humans in losses of food and industrial usage. Typical example is in the potato blight pandemic of 1845-1852 that affected many countries in Europe. Potato blight is a plant infection caused by the fungus *Phytopthera infestans*. About 1 million to 1.5 million people died while another 1 million to 1.5 million emigrated from Ireland to other countries (Powderly, 2019). Such massive loss can be prevented with prompt detection of fungal pathogens. The only option to prevent losses from root rot in peas is in understanding the disease, identifying the risks for infection and thorough planning. This study aimed at the possibility of detecting fungi and oomycetes pathogens of root rot of peas. Pea consumption is global, pea soup and pancakes is a tradition in Sweden that dates back to the Middle Ages. Root rot is the predominant disease in peas, food security would be improved by efforts geared at ensuring an efficient and fast detection of these pathogens.

Future perspectives

There are many other fungi that interfere with the identification of pea root rot specific pathogens (Manning et al., 1980). Also, the pea root and rootlets easily break off in infected tissues into the soil (Scott, 1987). Pathogens causing root rots are difficult to identify in severely damaged or dead plants because of the presence of other organisms that feed on the decaying tissue. Optimal results can be achieved with more attention or focus on the age and state of the samples. The age of plant has direct impact on the DNA extracted from such plants (Handayani et al., 2021). For the future studies, the samples should be analysed for integrity to improve the reliability and suitability of each sample.

The metagenomics workflow is seemingly simple but can be easily compromised based on the sensitivity of the different steps. Contaminants such as phenols and other phenolic structures of humic substances interferes with downstream applications of extracted DNA especially in PCR (Choi et al., 2015). Furthermore, the similar characteristics in size and charge of these phenolic substances affects DNA quantification (Albers et al., 2013). Cleaning of the extracted DNA to remove humic acid as demonstrated by Robe et al. (2003) with the use of Sepharose resins prior

quantification and downstream applications is important. This will be a worthwhile recommendation for future studies.

Acknowledgements

I sincerely appreciate the University of Skövde for the funds availability for this study and my thesis supervisor, Maria Algerin (Institution of Bioscience, University of Skövde) for her guidance and patient assistant. Also, of immense and timely assistance are Johan Nordén (co-supervisor), and John Baxter (for MinIon Sequencing and analysis). Special thanks to my examiner, Magnus Fagerlind.

My gratitude to my friends (The Johanssons, The Bengsts and ELIM kyrkan, Skövde, brethren). Also, I appreciate the love, sacrifice and support of my wife and children (Ibukun, Ini and Ire), God bless you all. All glory and adoration be to God in Jesus name!

References

Abarenkov, K., Zirk, A., Piirmann, Pöhönen, R., Ivanov, F., Nilsson, R. H., & Kõljalg, U. (2021): UNITE general FASTA release for eukaryotes. *UNITE Community*. doi.10.15156/BIO/1280160

Albers, C. N., Jensen, A., Bælum, J., & Jacobsen, C. S. (2013). Inhibition of DNA polymerases used in q-PCR by structurally different soil-derived humic substances. *Geomicrobiol J*. doi: 10.1080/01490451.2012.758193

Alberta Pulse Growers. (2014, December 22). *Root Rot in Pea and Lentil in Western Canada (PCN Winter 2015)*. https://albertapulse.com/2014/12/root-rot-pea-lentil-western-canada-pcn-winter-2015

Alford, D. V. (2008). *Biocontrol of Oilseed Rape Pests.* Blackwell publishing.

Arseneau, J. R., Steeves, R., & Laflamme, M. (2017). Modified low salt CTAB extraction of high quality DNA from contaminant rich tissues. *Molecular ecology resources*, 17(4), 686-693

Ashby, S. F., & Nowell, W. (1926). The Fungi of Stigmatomycosis. *Ann Bot*. os-40 (1): 69 –84. doi:10.1093/oxfordjournals.aob.a090018

Bainard, L. D., Navarro-Borrell, A., Hamel, C., Braun, K., Hanson, K. & Gan, Y. (2017). Increasing the frequency of pulses in crop rotations reduces soil fungal diversity and increases the proportion of fungal pathotrophs in a semiarid agroecosystem. *Agric. Ecosyst. Environ.* 240, 206–214. doi: 10.1016/j.agee.2017.02.020

Barnett, J. A. (1975). The entry of D-ribose into some yeasts of the genus Pichia. *Journal of General Microbiology*. 90 (1): 1–12. doi: 10.1099/00221287-90-1-1. PMID 1176959

Beakes, G. W. & Thines, M. (2016). Hyphochytriomycota and Oomycota. *Handbook of the Protists..* doi.org/10.1007/978-3-319-32669-6

Begerow, D., Nilsson, H., Unterseher, M. & Maier, W. (2010). Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Applied Microbiology and Biotechnology*, 87, pp. 99-108. doi.org/10.1007/s00253-010-2585-4

Bodah, E. T., Porter, L. D., Chaves, B., & Dhingra, A. (2016). Evaluation of pea accessions and commercial cultivars for fusarium root rot resistance. *Euphytica,* 208, 63–72 doi: 10.1007/s10681-015-1545-6

Bongam, A. L. (2022). Detection of soil borne pathogens causing pea root rot using Minion. *Master Degree project in Bioscience*, University of Skovde.

Botstein, D., & Fink, G. R. (2011). Yeast: an experimental organism for 21st Century biology. *Genetics*. 189 (3): 695–704. doi:10.1534/genetics.111.130765. PMC 3213361. PMID 22084421

Buée M., De Boer, W., Martin, F., Van Overbeek, L., & Jurkevitch, E. (2009). The rhizosphere zoo: An overview of plant associated communities of microorganisms, including phages, bacteria, archaea, fungi and of some of their structuring factors. *Plant and Soil*. 321 (1–2). 189–212

Chaparro, J. M., Badri, D. V., Bakker, M. G., Sugiyama, A., Manter, D. K., & Vivanco, J. M. (2013). Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally

programmed and correlate with soil microbial functions. *PloS One* 8, doi: 10.1371/journal.pone.0055731

Chatterton, S., Bowness, R., & Harding, M.W (2015). First Report of Root Rot of Field Pea Caused by Aphanomyces euteiches in Alberta, Canada. *Plant Dis*. doi: 10.1094/PDIS-09-14-0905-PDN. PMID: 30699595.

Chen, X., Jiang, Z. H., Chen, S., & Qin, W. (2010). Microbial and bioconversion production of D-xylitol and its detection and application. *International Journal of Biological Sciences*. 6 (7): 834–844. doi:10.7150/ijbs.6.834. PMC 3005349. PMID 21179590

Chiang, K. S., Liu, S. C., Bock, C. H. & Gottwald, T. R. (2017). A discussion on disease severity values. Part I. Warning on inherent errors and suggestions to maximise accuracy.*Annals of applied biology*. 171(2), 139-154

Choi J, Lee H, & Shipunov A. (2015). All that is gold does not glitter? Age, taxonomy, and ancient plant DNA quality. *Peer Life & Environment.* doi: 10.7717/peerj.1087. PMID: 26244108; PMCID: PMC4517958

Datto, M., & Lundbla, R. L. (2016). DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing). *Encyclopedia of Cell Biology*. Academic Press.

Deacon, J. W. (2005). *Fungal biology*. John Wiley & Sons.

Duke, J. A. (1981). *Hand book of legumes of world economic importance.* Plenum Press.

Eliane, T. B. (2017). Root Rot Diseases in Plants: A Review of Common Causal Agents and Management Strategies. *Agriculture Research and Technology Open Access Journal.* ISSN: 2471-6774

Euskirchen, P., Bielle, F., Labreche, K., Kloosterman, W. P., Rosenberg, S., Daniau, M., Schmitt, C., Masliah-Planchon, J., Bourdeaut, F., Dehais, C., Marie, Y., Delattre, J. Y., & Idbaih, A. (2017). Same day genomic and epigenomic diagnosis of brain tumors using realtime nanopore sequencing. *Acta Neuropathol.* doi: 10.1007/s00401-017-1743-5

Fink, H. C., & Buchholtz, W. F. (1954). Correlation between sugar beet crop losses and greenhouse determinations of soil infestation of Aphanomyces cochlioides. *Proc. Am. Soc. Sugar Beet Technol.* 8:252-259

Gardes, M., & Bruns, T.D. (1993). ITS primers with enhanced specificity for basidiomycetes: application to the identification of mycorrhiza and rusts. *Molecular Ecology.* doi: 10.1111/j.1365-294X.1993.tb00005

Gladieux, P., Wilson, B. A., Perraudeau, F., Montoya, L. A., Kowbel, D., Hann-Soden, C., Fischer, M., Sylvain, I., Jacobson, D. J., & Taylor, J. W. (2015) Genomic sequencing reveals historical, demographic and selective factors associated with the diversification of the fire-associated fungus Neurospora discreta. *Molecular Ecology*. doi.org/10.1111/mec.13417

Gossen, B. D., Conner, R. L., Chang, K. F., Pasche, J. S., McLaren, D. L. & Henriquez, M. A. (2016). Identifying and managing root rot of pulses on the Northern Great plains. *Plant Diseases*. doi: 10.1094/PDIS-02-16-0184

Handayani, E., Nandariyah, V. R., Cahyani & Parjanto (2021). The Effect of Weight and Incubation Time in DNA Quality of Kepel (Stelechocarpus burahol) Leaves. *IOP Conf. Ser.: Earth Environ. Sci*. doi:10.1088/1755-1315/752/1/012017

Hawksworth, D. L., & Lücking, R. (2017). Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *The Fungal Kingdom Microbiology Spectrum*. doi: 10.1128/microbiolspec.

Hyde, K. D., Xu, J. C., Lumyong, S., Rapior, S., Jeewon, R., Lumyong, S., Niego, G. T., Abeywickrama, P.D., Aluthmuhandiram, J.V., Brahamanage, R.S., Brooks, S., Chaiyasen, A., Chethana, K.W.T., Chomnunti, P., Chepkirui, C., Chuankid, B., de Silva, N.I., Doilom, M., Faulds, C., … Stadler, M. (2019). The amazing potential of fungi, 50 ways we can exploit fungi industrially. *Fungal Diversity*. doi.org/10.1007/s13225-019-00430-9

Infantino, A., Kharrat, M., Riccioni, L., Coyne, C. J., McPhee, K. E., & Grünwald, N. J. (2006). Screening techniques and sources of resistance to root diseases in cool season food legumes. *Euphytica. d*oi: 10.1007/s10681-006-6963-Z

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biology*. doi.org/10.1186/s13059-016-1103-0

Johnson, S., Maziade, P. J., McFarland, L. V., Trick, W., Donskey, C., Currie, B., Low, D. E., & Goldstein, E. J. (2012). "Is primary prevention of Clostridium difficile infection possible with specific probiotics?". *International Journal of Infectious Diseases*. doi:10.1016/j.ijid.2012.06.005

Kumari, N., & Katoch, S. (2020). Wilt and root rot complex of important pulse crops: their detection and integrated management. *Management of Fungal Pathogens in Pulses*. doi: 10.1007/978-3-030-35947-8

Li, X. G., Ding, C. F., Zhang, T. L., & Wang, X. X. (2014). Fungal pathogen accumulation at the expense of plant-beneficial fungi as a consequence of consecutive peanut monoculturing. *Soil Biol. Biochem*. doi: 10.1016/j.soilbio.2014.01.019

Liu, Z., Wang, Y., Deng, T., & Chen, Q. (2016). Solid-State Nanopore-Based DNA Sequencing Technology. *Journal of Nanomaterials*. 1−13.

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.* doi.org/10.1128/AEM .01368-19.

Lorenz, T. C. (2012) Polymerase chain reaction basic protocol plus troubleshooting and optimization strategies. *J Vis Exp*. doi: 10,3791/3998.

Lukas, W., Mario K., Monika, M., Bruno, S., & Pierre, H. (2021). Untangling the Pea Root Rot Complex Reveals Microbial Markers for Plant Health. *Frontiers in Plant Science.*

Lynch, J. P. (2007). Roots of the second green revolution. *Aust. J. Bot*.

Manning, M.A., Menzies, S.A. (1980). Root rot of peas in New Zealand caused by Aphanomyces euteiches N Z. *Agric Res*.

Martin, P. L., King, W., Bell, T. H., & Peter, K. (2021). The decay and fungal succession of apples with bitter rot across a vegetation diversity gradient. *Phytobiomes Journal.* doi: 10.1094/pbiomes-06-21-0039

Mendes, R., Garbeva, P. & Raaijmakers, J.M. (2013). The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms". *FEMS Microbiology Reviews*.

Naranjo, O., Miguel, A., Gabaldón, & Toni. (2019). Fungal evolution: Diversity, taxonomy and phylogeny of the Fungi. *Biological Reviews*. doi:10.1111/brv.12550

Niedringhaus, T. P., Milanova, D., Kerby, M. B., Snyder, M. P., & Barron, A. E. (2011). Landscape of next-generation sequencing technologies. *Analytical Chemistry*.

Nilsson, R. H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M., Porter, T. M., Bengtsson-Palme, J., Walker, D. M., de Sousa, F., Gamper, H. A., Larsson, E., Larsson, K. H., Kõljalg, U., Edgar, R. C., & Abarenkov, K. A. (2015). Comprehensive Automatically Updated Fungal ITS Sequence Dataset for Reference-Based Chimera Control in Environmental Sequencing Efforts. *Microbes Environ*. doi:10.1264/jsme2.ME14121.

Nilsson, R. H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., & Tedersoo, L. (2018). Mycobiome diversity: High-throughput sequencing and identification of fungi. *Nature Reviews in Microbiology*. doi.org/10.1038/s4157 9-018-0116

Nutter, F. W., Teng, P. S., & Shokes, F. M. (1991). Disease assessment terms and concepts. *Plant Disease.* 75.1187−1188

Ondov, B. D., Bergman, N. H., & Phillippy A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*. 12(1):385.

Peng, W., Kong, X., Chen, Y., Zhang, C., Yang, Y., & Hua, Y. (2016). Effects of heat treatment on the emulsifying properties of pea proteins. *Food Hydrocoll.* 52, 301−310. doi: 10.1016/j.foodhyd.2015.06.025

Powderly, W. G. (2019). How infection shaped history: Lessons from the Irish famine. *Trans Am Clin Climatol Assoc*.

Rantala-Ylinen, A., Känä, S., Wang, H., Rouhiainen, L., Wahlsten, M., Rizzi, E., Berg, K., Gugger, M. & Sivonen, K. (2011). Anatoxin-a synthetase gene cluster of the cyanobacterium Anabaena sp. strain 37 and molecular methods to detect potential producers. *Appl Environ Microbiol.* doi: 10.1128/AEM.06022-11.

Rubiales, D., Fondevilla, S., Chen, W., Gentzbittel, L., Higgins, T. J., & Castillejo, M. A., (2015). Achievements and challenges in legume breeding for pest and disease resistance. *Crit. Rev. Plant Sci.* doi: 10.1080/07352689

Robe, P., Nalin, R., Capellano, C., Vogel, T. M., & Simonet, P. (2003). Extraction of DNA from soil. *European Journal of Soil Biology*

Rogers, S. (2007). *Man and the Biological World*. Read Books

Sasse, J., Martinoia, E., Northen, T. (2017). Feed your friends: Do plant exudates shape the root microbiome? *Trends Plant Sci*. doi: 10.1016/j.tplants

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* doi:10,1073/pnas.74.12.5463

Sarah, C., Castle, Zewei S., Daryl, M., Gohl, Jessica, L., Gutknecht, M., Carl, J., Rosen, Michael, J., Sadowsky, Deborah, A., Samac, & Linda, L. K. (2018). DNA Template Dilution Impacts Amplicon Sequencing-Based Estimates of Soil Fungal Diversity. doi.org/10.1094/PBIOMES

Scott, R.E. (1987). Root rot and soil compaction problems of pea crops. *Agron Soc. New Zealand Spec Publ*.

Schalamun, M., Nagar, R., Kainer, D., Beavan, E., Eccles, D., Rathjen, J. P., Lanfear, R., & Schwessinger, B. (2019). Harnessing the MinION: An example of how to establish long read sequencing in a laboratory using challenging plant tissue from Eucalyptus pauciflora. *Molecular ecology resources*

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., & Chen, W. (2012). Fungal Barcoding Consortium; Fungal Barcoding Consortium Author List. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci* doi: 10.1073/pnas.1117018109.

Sharma, K., Mishra, A. K., & Misra, R. S. (2008). A simple and efficient method for extraction of genomic DNA from tropical tuber crops. *African Journal of biotechnology.*

Sidstedt, M., Rådström, P. & Hedman, J. (2020). PCR inhibition in qPCR, dPCR, MPS-mechanisms and solutions. *Anal Bioanal Chem*. doi: 10.1007/s00216-020-02490-2

Sipos, R., Székely, A.J., Palatinszky, M., Révész, S., Márialigeti, K., Nikolausz, M. (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiology Ecology*. doi: 10.1111/j.1574-6941

Smith, D.J., Yadav, T., & Whitehouse, I. (2015). Detection and Sequencing of Okazaki Fragments in *S. cerevisiae*. *Methods Mol Biol.* doi: 10.1007/978-1-4939-2596-4_10

Smith, R.L., Sawbridge, T., Mann, R., Kaur, J., May, T.W., & Edwards, J. (2020). Rediscovering an old foe: Optimised molecular methods for DNA extraction and sequencing applications for fungarium specimens of powdery mildew (Erysiphales). *PLoS One*. doi: 10.1371/journal.pone

Sutlovic, D., Gamulin, S., Definis-Gojanovic, M., Gugic, D., Andjelinovic, S.(2008). Interaction of humic acids with human DNA: Proposed mechanisms and kinetics. *Electrophoresis.* doi: 10.1002/elps.200700699.

Tebbe, C. C., & Vahjen, W. (1993). Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant-DNA from bacteria and a yeast. *Appl Environ Microbio*l. doi: 10.1128/AEM.59.8.2657-2665.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N, S., Wijesundera, R., Villarreal, R. L., Vasco-Palacios, A, M., Thu, P, Q., Suija, A., Smith, M, E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa … Abarenkov, K. (2014). Global diversity and geography of soil fungi. *Science* doi: 10.1126/science.1256688

Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R.H., Kennedy, P.G., Yang, T., Anslan, S., & Mikryukov, V. (2022).Best practices in metabarcoding of fungi: From experimental design to results. *Mol Ecol.* doi: 10.1111/mec.16460

Tsai, Y. L., & Olson, B. H. (1992). Detection of low numbers of bacterial numbers in soils and sediments by polymerase chain reaction. *Applied and Environmental Microbiology*. doi: 10.1128/aem

Wallenhammar, A. C., Almquist, C., Söderström, M., & Jonsson, A. (2012). In field distribution of Plasmodiophora brassicae measured using quantitative real- time PCR. *Plant Pathology*

Wibberg, D,, Genzel, F., Verwaaijen, B., Blom, J., Rupp, O., Goesmann, A., Zrenner, R., Grosch, R., Pühler, A. & Schlüter, A. (2021). Genome Analyses of the Less Aggressive Rhizoctonia solani AG1-IB Isolates 1/2/21 and O8/2 Compared to the Reference AG1-IB Isolate 7/3/14. *J Fungi (Basel)*. doi: 10.3390/jof7100832.

Wilfinger, W. W., Mackey, K., & Chomczynski, P. (1997). Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques*

Wood, D. E. & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*

Wood, D.E., Lu, J. & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology.* doi.org/10.1186/s13059-019-1891-0

Wurzbacher, C., Larsson, E., Bengtsson-Palme, J., Van den Wyngaert, S., Svantesson, S., & Kristiansson, E. (2019). Introducing ribosomal tandem repeat barcoding for fungi. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12944

Xiong, T., Xiong, W., Ge, M., Xia, J., Li, B. & Chen, Y. (2018). Effect of high intensity ultrasound on structure and foaming properties of pea protein isolate. *Food Res. Int*. doi: 10.1016/j.foodres

Xu, L., Ravnskov, S., Larsen, J., Nilsson, R. H., & Nicolaisen, M. (2012). Soil fungal community structure along a soil health gradient in pea fields examined using deep amplicon sequencing. *Soil Biol. Biochem*. doi: 10.1016/j.

Zhu, T. Y., Wang, B., & Guo, Q. (2006). Isolate high quality DNA from varieties of plant specimens with E.Z.N.A.TM plant DNA systems from Omega Bio-Tek. *VWR bioMarke.*

Appendix

Appendix 1- Primer sequences analysed using Oligo Analyser and illustration of expected ITS regions

Primer sequences analysed using Oligo Analyser.

Primers	Sequence 5'-3'	length	[a]GC	[b]MT	[c]MW	*Hairpin Kcal/mol	*Dimers Kcal/mo
ITS1OO	GGAAGGATCATTACCACAC	19	47	50	5805	-0.20	-4.62
ITS1Catta	ACCWGCGGARGGATCATTA	19	50	53	5849	-0.19	-5.40
ITS4ngs	TCCTSCGCTTATTGATATGC	20	45	52	6054	0.30	-9.69
ITS4ngsUni	CCTSCSCTTANTDATATGC	19	47	47	5720	2.39	-9.69

*Hairpins and *dimers are formed as secondary structures when mRNA strands folds form base pairs with another section of the strand, Delta G values in Kcal/mol denotes the stability of the structures formed. (ΔG= -10 Kcal/mol and less is usually tolerated).

The Secondary structures and annealing temperatures of the primers pairs calculated with Oligo Analyser.

[a]GC GC content
[b]MT Melting Temperature
[c]MW Molecular Weight

Primer pairs features.

Primer pair	Annealing temperature (°C)	Hetero dimers (Kcal/mol)
ITS1catta-ITS4ngsUni	63.20	-17.58
ITS1OO-ITS4ngs	62.00	-4.64

Fusarium solani and *Phytophthora vignae* ITS region (NCBI), showing forward and reverse primer. Expected amplicon base pairs highlighted in grey

>NR_163531.1 Fusarium solani CBS 140079 ITS region; from TYPE material

Forward primer: ITS1Catta

5'ACCAGCGGAGGGATCATTA3'
→

5'TGGAAGTAAAAAGTCGTAACAAGGTCTCCGTTGGTGAACCAGCGGAGGGATCATTACCGAGTTATA
CAACTCATCAACCCTGTGAACATACCTAAAACGTTGCTTCGGCGGGAACAGACGGCCCTGTAACAACG
GGCCGCCCCCGCCAGAGGACCCCTAACTCTGTTTTTATAATGTTTTTCTGAGTAAACAAGCAAATAAA
TTAAAACTTTCAACAACGGATCTCTTGGCTCTGGCATCGATGAAGAACGCAGCGAAATGCGATAAGTA
ATGTGAATTGCAGAATTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCGCCAGTATTCTGGC
GGGCATGCCTGTTCGAGCGTCATTACAACCCTCAGGCCCCCGGGCCTGGCGTTGGGGATCGGCAGAAG
CCCCCTGTGGGCACACGCCGTCCCTCAAATACAGTGGCGGTCCCGCCGCAGCTTCCATTGCGTAGTAG
CTAACACCTCGCAACTGGAGAGCGGCGCGGCCATGCCGTAAAACACCCAACTTCTGAATGTTGACCTC
GAATCAGGTAGGAATACCCGCTGAACTTAAGCATATCAATAAGCGGAGGAAAAGAAACCAACAGGGAT
TGCCCCAGTAACGGCGAGTGAA 3'

CGTATAGTTATTCGCCTCCT
←
Reverse primer: ITS4ngsUni

>MW476179.1 *Phytophthora vignae* isolate

Forward primer: ITS100

5′GGAAGGATCATTACCACAC
→

5′TTCCGTAGGTGAACCTGCGGAAGGATCATTACCACACCTAAAAAACTTTCCACGTGAACCGTTTCA
ACAAGTAGTTGGGGGCCTGCTCTGTGTGGCTAGCTGTCGATGTCAAAGTCGGCGACTGGCTGCTATGT
GACGGGCTCTATCATGGCAATTGGTTTGGGTCCTCCTCGTGGGGAACTAGATCATGAGCCCACTTTTT
AAACCCATTCTTGATTACTGAATATACTGTGGGGACGAAAGTCTCTGCTTTTAACTAGATAGCAACTT
TCAGCAGTGGATGTCTAGGCTCGCACATCGATGAAGAACGCTGCGAACTGCGATACGTAATGCGAATT
GCAGGATTCAGTGAGTCATCGAAATTTTGAACGCATATTGCACTTCCGGGTTAGTCCTGGGAGTATGC
CTGTATCAGTGTCCGTACATCAAACTTGGCTCTCTTCCTTCCGTGTAGTCGGTGGATGGAGACGCCAG
ACGTGAGGTGTCTTGCGGCGCGGCCTTCGGGCTGCCTGCGAGTCCCTTGAAATGTACTGAACTGTACT
TCTCTTTGCTCGAAAAGCGTGACGTTGTTGGTTGTGGAGGCTGCCTGTATGGCCAGTCGGCGACCAGT
TTGTCTGCTGCGGCGTTTAATGGAGGAGTGTTCGATTCGCGGTATGGTTGGCTTCGGCTGAACAATGC
GCTTATTGGATGCTTTTCCTGCTGTGGCGGTATGGGCTGGTGAACCGTAGTTGTGCAAGGCTTGGCTT
TTGAACCGGTGGTGTTGTTGCGAAGTAGGGTGGCGGCTTCGGCTGTCGAGGGTCGATCCATTTGGGAA
CTTTGTGTTGTCTCTGCGGCTTGCTGTGGAG 3′

5′TCCTSCGCTTATTGATATGC 3′
Reverse primer: ITS4ngs
←

Appendix 2- Samples for sequencing

Sample	DNA Type	Primer pairs
15020	Amplicon	ITS1catta & ITS4ngsUni
ORSMARK2	Amplicon	ITS1catta & ITS4ngsUni
14735	Amplicon	ITS1catta & ITS4ngsUni
15020	Amplicon	ITS100 & ITS4ngs
ORSMARK2	Amplicon	ITS100 & ITS4ngs
14735	Amplicon	ITS100 & ITS4ngs
15020	Genomic	-
ORSMARK2	Genomic	
14735	Genomic	

Appendix 3- Concentration of samples after end preparation and native barcoding

Qubit concentration of samples after end preparation and native barcoding ligation steps

Primer pair	Sample ID	Barcode	End preparation Concentration (ng/µl)	Native barcoding concentration (ng/µl)
A	15020	13	0.68	0.19
A	ORSMARK2	14	1.07	0.74
A	14735	15	0.88	0.67
B	15020	17	0.68	0.25
B	ORSMARK2	18	1.45	1.37
B	14735	19	1.26	0.95
Genomic DNA	15020	21	1.15	0.77
Genomic DNA	ORSMARK2	22	1.08	0.85
Genomic DNA	14735	23	0.88	0.67

* **A- ITS1catta & ITS4ngsUni, B- ITS1OO & ITS4ngs**

Appendix 4 - Qubit concentration of barcode amplicon and DNA amount

Qubit concentration of barcode amplicon and the amount of DNA used for sequencing.

Primer pair	Sample ID	Barcode	Concentration of barcode samples (ng/µl)	Quantity used (µl)
A	15020	13	0.19	6.60
A	ORSMARK2	14	0.74	2.72
A	14735	15	0.67	3.00
B	15020	17	0.25	8.04
B	ORSMARK2	18	1.37	1.47
B	14735	19	0.95	2.12
Unamplified DNA	15020	21	0.56	3.59
Unamplified DNA	ORSMARK2	22	0.77	2.61
Unamplified DNA	14735	23	0.85	2.36

* **A- ITS1catta & ITS4ngsUni, B- ITS1OO & ITS4ngs**

Appendix 5 - Sequencing performance of the samples.

Sequencing performance of the samples.

Sample	DNA Type	Primer pairs	DSI	Classified Reads (%)		Unclassified Reads (%)	
				Kraken2	UNITE	Kraken2	UNITE
15020	Amplicon	A	66	4.9	9.6	95.1	90.4
ORSMARK2	Amplicon	A	36	1.6	5.9	98.4	94.1
14735	Amplicon	A	28	1.0	4.2	99	95.8
15020	Amplicon	B	66	2.3	8.9	97.7	91.1
ORSMARK2	Amplicon	B	36	1.7	5.6	98.3	94.4
14735	Amplicon	B	28	1.3	5.4	98.7	94.6
15020	Genomic	-	66	1.3	4.4	98.7	95.6
ORSMARK2	Genomic	-	36	1.1	3.7	98.9	96.3
14735	Genomic	-	28	1.0	3.2	99.0	96.8

* **A- ITS1catta & ITS4ngsUni, B- ITS1OO & ITS4ngs**