Population Analysis and Subgroup Diagnosis of

Lettuce necrotic yellows virus

within New Zealand

Toni Louise Darling

A thesis submitted to
Auckland University of Technology
In partial fulfilment of the requirements for the degree of
Master of Science (Research)
MSc (Research)

2020

School of Science

Abstract

Lettuce is a key economic crop in New Zealand (NZ). However, recently NZ has experienced increased crop losses of up to 50 %. These losses were due to necrosis symptoms on leaves, which were strongly associated with the plant virus lettuce necrotic yellows virus (LNYV). This virus has also been found in Australia where heavy lettuce crop losses have also occurred since 1954.

Lettuce necrotic yellows virus is the type species of the genus *Cytorhabdovirus*, family *Rhabdoviridae*, commonly referred to as rhabdoviruses. This negative sense, single stranded RNA plant virus can be differentiated into two subgroups (I and II) based on genetic differences within its nucleoprotein (N).

The LNYV subgroup population structure has shown to be dissimilar between NZ and Australia. In Australia, no new isolates belonging to subgroup I have been identified since 1993. It appears this subgroup has become extinct, possibly due to subgroup II. It has been suggested subgroup II may have supplanted subgroup I due to more optimal vector and/or host interactions. However, in NZ, both subgroups are still being identified across the country. It has been suggested that the high crop losses experienced in NZ recently could be due to an increased presence of subgroup II, or a more virulent strain of the virus has arrived here from Australia.

To investigate the cause of these recent losses in NZ, a closed tube, quick, sensitive and specific diagnostic assay is required. The two LNYV subgroups have not demonstrated discriminating symptoms or detectable serological differences. This has meant subgroup identification needs to occur on a molecular level. In 2018, a molecular assay to diagnose the LNYV subgroups was developed using reverse transcription polymerase chain reaction (RT-PCR). Unfortunately, the developed assay had limitations, requiring further analysis using restriction fragment length polymorphism, or for each sample to be tested twice – once for each subgroup – for confidence in the diagnosis of LNYV infection and subgroup identification. It was suggested the developed LNYV subgroup specific primers could be suitable for use in an alternative molecular assay - multiplex reverse transcription quantitative polymerase chain reaction with high resolution melting (RT-qPCR-HRM).

This study has focused on the development of such an assay. First the LNYV subgroup specific primers from the previous assay were tested for their suitability in a multiplex RT-qPCR-HRM assay. Through this analysis several qPCR fluorescent dyes were assessed. Although SYBR Green fluorescent dye is commonly used, several studies have shown it to inhibit PCR and the

dye molecules to translocate during HRM analysis, affecting the accuracy of the assay. This was also indicated during this study, and eventually the BioRad Sso Fast EvaGreen dye was chosen as it demonstrated superior sensitivity.

In this study it was discovered that the LNYV subgroup specific primers were unable to sufficiently discriminate between the two subgroups using HRM. This meant LNYV subgroup specific primers designed specifically for use in multiplex RT-qPCR-HRM analysis needed to be developed. Three primer sets were designed and tested for their usefulness in singleplex and multiplex assays. Two primer sets appeared suitable, so were used to diagnose the LNYV subgroup of previously untyped LNYV isolates. Through this analysis, one primer set appeared able to detect the LNYV subgroup, while also distinguishing variances in the amplified PCR product sequences. These variances fit a quasi-species model, which is common among RNA viruses. However, this is the first suggestion that LNYV exists as a quasi-species. This analysis also identified, for the first time, three samples co-infected with both LNYV subgroups.

The LNYV subgroup specific RT-qPCR-HRM diagnostic assay developed in this study identified subgroup I and subgroup II samples from both the North and South Islands of NZ. However, of the 18 untyped samples, only four were identified as subgroup II, with a further three identified as co-infected. This indicated that subgroup II was not exhibiting a stronger presence in NZ at this time. To assess whether a new LNYV strain from Australia had arrived in NZ additional further analysis of these isolates was required.

The entire N gene from these isolates was sequenced to characterise their taxonomic and evolutionary relationships using phylogenetics. A total of 43 LNYV isolates, consisting of published and unpublished sequences, from NZ and Australia were used to assess these relationships. This analysis was unable to identify an isolate of direct Australian lineage that would indicate a stronger strain had arrived in NZ. However, the analysis did reveal other relationships. LNYV subgroup II isolates from both Australia and NZ appeared to originate in Australia, while the origin for subgroup I isolates was unclear. However, within the NZ subgroup I isolates, geographic specific groupings were observed.

Keywords: Lettuce necrotic yellows virus, subgroup, RT-qPCR, high resolution melting, phylogenetics

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signed:

Toni Louise Darling

Date:

10 August 2020

Acknowledgements

I sincerely wish to thank all those who have supported me and therefore contributed in one way or another to the completion of this thesis. You are all awesome people who deserve recognition for this.

Firstly, to Dr. Colleen Higgins; supervisor and wonder woman, it has been a pleasure to perform research under your guidance and extensive expertise once again. My heartfelt thanks for taking me on board, reawakening my passion for research and for your continued support, knowledge and leadership. Through this project you have tried to teach me self-discipline, while still being flexible, approachable and fair. I am so thankful to have had you as a role model for research, career, and life.

To the AUT Molecular Genetics Research Group, my deepest gratitude for making the weekly meetings more than just sharing experimental results. Thank you for your support, encouragements, reading my chapters and laughs! Special thanks to Anthony Hull and Elizabeth Buckley for sharing your knowledge and showing me the ropes around the lab, to Shweta Shinde and Lee Rabbidge for your support during my first overseas conference, and to Zoe King for your inspiration and calmness during the writing phase.

I would also like to thank John Fletcher, formerly of The Institute for Plant & Food Research, for the supply of LNYV infected lettuce from around NZ, and the Australasian Plant Pathology Society for their funding and support to attend their conference and present at their meeting.

My appreciation goes to those amazing friends outside of the university who offered support through asking, "how's the writing going?" and listening through my ups and downs; Gina Swart, Sunny Jamati, Deanna Robertson, Danielle Carter, Cole Beer and May Alfante.

I can't begin to thank all my family, the Darling and the Davis, enough for their consistent support, encouragement and love. There were plenty of laughs when I got messages such as "what are you studying again?" and "Toni can't do that; she's got a thesis to write!" A very special mention to Mum, Debbie Darling, you've always encouraged me with my career and further education and your support through this time, as I said 'yes' to too many things, has been irreplaceable. Likewise, Ross Davis, thank you for your willingness to support my continued study at this stage of our lives and pushing me to do well.

List of Abbreviations

+G	Gamma distribution
+I	Evolutionarily invariable
$	Dollars
°C	Degree celsius
µg/mL	Microgram per millilitre
µL	Microlitre
µM	Micromolar
2-ME	2-mercaptoethanol
4b	LNYV MP
ADV	*Alfalfa dwarf virus*
ArMV	*Arabis mosaic virus*
BASTA	Bayesian Structured Coalescent Approximation
BEAST	Bayesian Evolutionary Analysis Sampling Trees
BIC	Bayesian Information Criterion
BioGeoBEARS	BioGeography with Bayesian Evolutionary Analysis in R Scripts
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
CBDaV	*Colocasia bobone disease-associated virus*
cDNA	complementary DNA
CGRMV	*Cherry green ring mottle virus*
CNRMV	*Cherry necrotic rusty mottle virus*
CoRSV	*Coffee ringspot virus*
DAS-ELISA	Double antibody sandwich ELISA
ddNTPs	Dideoxynucleotide triphosphates
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
dsDNA	Double-stranded DNA
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
EMDV	*Eggplant mottled dwarf virus*
EtBr	Ethidium bromide
G	Glycoprotein
GFLV	*Grapevine fanleaf virus*
GLD	*Grapevine leafroll disease*
GLRaV-1	*Grapevine leafroll-associated virus 1*
GLRaV-3	*Grapevine leafroll-associated virus 3*
ha	Hectares
HIV	*Human immunodeficiency virus*
HRM	High resolution melting
HSVd	*hop stunt viroid*
HTS	high-throughput sequencing
ICTV	International Committee on Taxonomy of Viruses
IDT	Integrated DNA Technologies
IHNv	*Infectious hematopoietic necrosis virus*
IUPAC	International Union of Pure and Applied Chemistry

IvLV	*Ivy latent virus*
JTT	Jones-Taylor-Thornton
kb	Kilobase
kg	Kilogram
L	RNA dependent RNA polymerase
LAMP	Loop-mediated isothermal amplification
LNYV	*Lettuce necrotic yellows virus*
LYMoV	*Lettuce yellow mottle virus*
M	Matrix
M	Million
MCA	Melt curve analysis
mg	Milligram
ML	Maximum likelihood
mL	Millilitre
MP	Movement protein
mRNA	Messenger RNA
N	Nucleoprotein
NCBI	National Centre for Biotechnology Information
NCMV	*Northern cereal mosaic virus*
NCPA	Nested Clade Phylogeographic Analysis
ng	Nanogram
NGS	Next generation sequencing
NJ	Neighbour-Joining
nm	Nanometre
nt	Nucleotides
NTC	No template control
NZ	New Zealand
OFV	*Orchid fleck virus*
ORFs	Open reading frames
P	Phosphoprotein
PCR	Polymerase chain reaction
PeVA	*Persimmon virus A*
PLRV	*Potato leafroll virus*
poly A	Polyadenylated
PPV	*Plum pox virus*
PVS	*Potato virus S*
PVX	*Potato virus X*
PVY	*Potato virus Y*
qPCR	Quantitative PCR
RCF	Relative centrifugal force
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
rRNA	Ribosomal RNA
RRV	*Rose rosette virus*
RT-PCR	Reverse transcription PCR
RT-qPCR	Reverse transcription quantitative PCR
RVCV	*Raspberry vein chlorosis virus*
SBS	Sequencing by synthesis
SCV	*Strawberry crinkle virus*

SDA	Strand displacement amplification
SNPs	Single nucleotide polymorphisms
SPREAD	Spatial Phylogenetics Reconstruction of Evolutionary Dynamics
SpreaD3	Spatial Phylogenetics Reconstruction of Evolutionary Dynamics using Data-Driven Documents
SSIV	SuperScript IV
T92	Tamura 3-parameter
TaVCV	*Taro vein chlorosis virus*
TBE	Tris/Borate/EDTA
TEM	Transmission electron microscopy
TI	Transcription initiation
T_m	Melting temperature
TTP	Transcription termination polyadenylation
UK	United Kingdom
UTR	Untranslated regions
VHSv	Viral haemorrhagic septicaemia virus

Chapter 1

General Introduction and Literature Review

1.1 General Introduction

The world's population is growing. By 2050 it is predicted to reach between 8.8 and 10 billion people (Spiertz 2010; Cleland 2013; Culliney 2014; Sharma 2014; Rao et al. 2016; Sharma et al. 2017). Consequently, global food requirements will also increase. An estimated increase of 70 % in global food production is needed to meet these demands (Culliney 2014; Fang and Ramasamy 2015; Rao et al. 2016; Sharma et al. 2017). Currently, agricultural systems occupy 38 % of Earth's total land area, this includes pasture and livestock grazing systems (Spiertz 2010; Culliney 2014; Savary et al. 2014). Previously, agricultural intensification – aiming for higher crop yields per unit of land – ensured the balance between population growth and food production (Savary et al. 2014). This is reflected by a drop in cropland area in relation to population; at the middle of the 20th century there was 0.45 hectares (ha) of cropland per person, but by 1997 there was only 0.25 ha per person, and by 2050 it is predicted to drop to 0.15 ha per person (Spiertz 2010; Sharma 2014).

The performance of agricultural systems can be enhanced by increasing the potential crop yields, the attainable crop yields, and/or the actual crop yield (Savary et al. 2014; Sharma et al. 2017). The potential crop yield is determined by growth-defining factors, such as temperature, UV radiation and the genetic makeup of the crop. This is the yield obtained when crops are grown under optimal conditions, maximising crop yield (Culliney 2014; Savary et al. 2014; Sharma et al. 2017). The attainable crop yield is achieved by growth-limiting factors, such as soil condition, nutrients and water supply. This is the site-specific, technical maximum yield. However, this yield is generally well below the potential crop yield (Culliney 2014; Savary et al. 2014; Sharma et al. 2017).

The actual crop yield is defined as the site-specific, real-world, level of production, which can be obtained within modern agricultural systems. It is affected by both growth-defining and growth-limiting factors, as well as growth-reducing factors, such as weeds, plant pathogens, insects and microclimate incidents such as frost or flooding (Culliney 2014; Savary et al. 2014; Sharma et al. 2017). Improvement of the actual crop yield addresses the quality as well as the quantity of the crop. It is also likely to produce yield improvements to the same degree as could be achieved by increasing the potential or attainable crop yields (Savary et al. 2014). Increasing the actual crop yield can be achieved by improving crop health, through the development of crop protection methods, and in particular, crop protection from harmful organisms (Savary et al. 2014; Sharma 2014).

Modern agricultural intensification has caused a rise in crop losses to plant pests and harmful organisms. Such practices entice plant pests due to monoculture and less plant diversity in fields, destruction of beneficial arthropods, and promotion of pesticide resistance, causing a change in the resident insect community (Anderson et al. 2003; Culliney 2014; Savary et al. 2014; Sharma 2014; Sharma et al. 2017). The resident insect community is therefore less diverse, when compared with traditional agriculture or natural ecosystems, but more abundant in the numbers of insects with a narrow host range relating to the available crops (Anderson et al. 2003; Culliney 2014).

Harmful organisms, or plant pests, include weeds, insects and other arthropods, diseases, and viruses (Dal Santo and Velthuis 2008; Savary et al. 2014; Fang and Ramasamy 2015). All insects (aphids, mealybugs, scale insects, spiders, flies, ticks and mites) are arthropods, and make up the most abundant, diverse and ubiquitous group of organisms on the planet (Culliney 2014; Sharma 2014; Sharma et al. 2017). Arthropods in particular can cause damage to crops in various ways. Some arthropods cause plant damage by chewing or sucking the sap from fruit, bark, stems, buds or leaves. They can also tunnel or bore into the plant and live in cancer-like growths. From here they can attack underground stems or root systems and lay eggs. Arthropods may also take parts of the plant away for nests or shelters (Culliney 2014).

Very recent data on crop losses is limited, although, it is estimated that direct yield losses due to plant pests ranges from 20 % – 40 % (Savary et al. 2014; Sharma et al. 2017). It has been estimated that arthropods, in particular, can cause 13 % – 16 % pre-harvest crop losses, and a further 5 % – 10 % post-harvest (Culliney 2014; Sharma 2014; Sharma et al. 2017). Furthermore, some arthropods cause indirect damage by just probing for their host plant, and thereby transmitting plant diseases and viruses (Stufkens et al. 2002; Redinbaugh and Hogenhout 2005; Diaz et al. 2012; Fang and Ramasamy 2015; Fletcher et al. 2019).

Globally, plant pathogen infections cause between 20 % and 40 % of crop yield losses (Anderson et al. 2003; Fang and Ramasamy 2015). After fungi, viruses are ranked as the most important plant pathogens, with economic losses in the billions of dollars per year, worldwide (Jeong et al. 2014; Fang and Ramasamy 2015; Pallás et al. 2018). In fact, a review of emerging infectious diseases of plants found viruses caused 47 % of plant diseases reported (Figure 1.1; Anderson et al. 2003). This is similar to that for human (44 %) and wildlife (~ 43 %) emerging infectious diseases at the time of publication (Anderson et al. 2003).
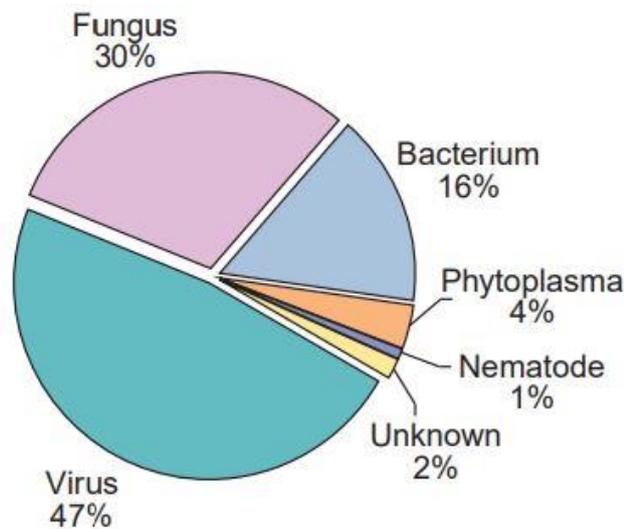
Figure 1.1 Summary of pathogens causing emerging infectious diseases in plants (Reproduced from Anderson et al. 2004).

1.1.1 The New Zealand Situation

The global circumstances are also reflected in New Zealand (NZ), with NZ surpassing global population growth rates of 1.1 % per year. In fact, NZ has exceeded this population growth rate, each year, over the last five years (StatsNZ 2019). NZ's population reached 4.92 million at the end of June 2019, an increase of 1.6 % (StatsNZ 2019). A report in 2017 from Horticulture NZ, stated that with current population growth, consumption, and production levels, NZ could expect food shortages in the next five years (Horticulture New Zealand [HortNZ] 2017). In 2016, NZ produced 1,133,800 tonnes of ten key vegetables (HortNZ 2017). While in 2018, NZ needed to import 118,185 tonnes of vegetables, with a value of NZ $ 263.8 million (M) from 80 different countries (Plant & Food Research [PFR] 2019).

In 2016, NZ household consumers spent NZ $ 930 M on fresh and chilled vegetables, and a further NZ $ 330 M on processed vegetables. The top three vegetables were potatoes and potato products, tomatoes and lettuce (PFR 2019). On average, New Zealanders consumed approximately 22.90 kgs of potatoes, 7.56 kgs of tomatoes, and 5.06 kgs of fresh lettuce, per person per year (HortNZ 2017). In that same year, NZ produced 525,000 tonnes of potatoes, 102,900 tonnes of tomatoes, and 8,400 tonnes of lettuce. However, a further 18.5 tonnes, 132.2 tonnes, and 122.5 tonnes of these vegetables, respectively, were imported to meet NZ's current supply demands (HortNZ 2017). The report issued by Horticulture NZ highlighted the importance of suitable land production, future-proofing the availability of resources, and food security, including biosecurity, of NZ's produce (HortNZ 2017).

NZ is one of the few countries (Fox and Mumford 2017) to keep a detailed database of plant pathogen species found in the country (Veerakone et al. 2015). As of the last update, 220 viruses, seven viroids, two liberibacters, and two phytoplasmas have been recorded (Veerakone et al. 2015). From the report potato (*Solanum tuberosum*) was identified as host to 11 viruses, one liberibacter, and one phytoplasma, tomato (*S. lycopersicum*) to 11 viruses and one liberibacter, and lettuce (*Lactuca sativa*) to ten viruses (Veerakone et al. 2015). This record provides information to NZ biosecurity departments so they are able to make informed regulations on import and export to and from NZ (Veerakone et al. 2015). The report also saw the addition of 80 viruses, one viroid, and two species of liberibacter as new to NZ since 2006 (Veerakone et al. 2015), emphasising the importance of biosecurity for NZ's produce.

Impacting NZ's production of lettuce, NZ lettuce growers have experienced crop losses of up to 50 % in recent years (Fletcher et al. 2019). Growers from Mid Canterbury, Nelson and the lower North Island saw losses in outdoor iceberg/crisp head varieties of lettuce. Crop losses due to necrosis symptoms on leaves were strongly associated with the plant virus lettuce necrotic yellows virus (LNYV; Fletcher et al. 2018; Fletcher et al. 2019). *Lettuce necrotic yellows virus* is the type species belonging to the genus *Cytorhabdovirus*, family *Rhabdoviridae*, commonly referred to as rhabdoviruses, and belonging to the order *Mononegavirales* (Dietzgen et al. 2007; Walker et al. 2018).

1.2 Literature Review

1.2.1 *Rhabdoviridae*

The *Rhabdoviridae* family consists of a diverse range of pathogens. These are classified across 22 genera with 144 species (Walker et al. 2018; Dietzgen et al. 2020). Viruses are assigned to each genus based on differences in genome architecture, antigenicity, ecological and biological properties, and phylogenetic analysis of the RNA-dependent RNA polymerase (L) protein sequences (Walker et al. 2018). Phylogenetic analysis, a bioinformatic tool, helps to organise and analyse inter-species and intra-species viral relationships (Section 1.2.3.3; Lam et al. 2010), allowing taxonomic and evolutionary relationships to be characterised. Using this bioinformatic analysis, members of *Rhabdoviridae* form monophyletic clades representing the different genera (Figure 1.2; Walker et al. 2018).

Rhabdoviruses can infect a large range of organisms including vertebrates, invertebrates, and plants. Many of these infections can have considerable impact on public health, veterinary, and/or agricultural production. The majority of rhabdovirus genera are transmitted by

arthropods to vertebrate or plant hosts (Figure 1.2; Walker et al. 2018). Exceptions to this include genera that infect vertebrate hosts but are not transmitted by an arthropod and genera that infect invertebrate hosts. Lyssaviruses are transmitted directly by bites or scratches to vertebrate hosts, tupaviruses have been isolated from birds, insectivores and rodents, while novirhabdovirus, perhabdovirus, and sprivivirus infect fish and are transmitted horizontally in the water. Almendraviruses were isolated from mosquitoes, alphanemrhavirsues were detected by high-throughput sequencing of parasitic nematodes but have not been isolated, caligrhaviruses, were detected in sea lice but have also not been isolated, and sigmaviruses, infect flies from the *Drosophilidae* or *Muscidae* families, and only transmit vertically across the specific fly species (Jackson et al. 2005; Redinbaugh and Hogenhout 2005; Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018; Whitfield et al. 2018).

Very recently the genus *Nucleorhabdovirus* has been further defined and is now split into three genera, *Alphanuceorhabdovirus, Betanucleorhabdovirus* and *Gammanucleorhabdovirus* (Dietzgen et al. 2020). Phylogenetic analyses of viruses belonging to these genera indicate paraphyletic origin since Datura yellow vein virus, Sonchus yellow net virus and Maize fine streak virus do not share the most recent common ancestor of the other nucleorhabdoviruses (Figure 1.2). This emphasises the diversity of virus species belonging to *Rhabdoviridae*, and the need for continual study to further develop our understanding and knowledge of these viruses.
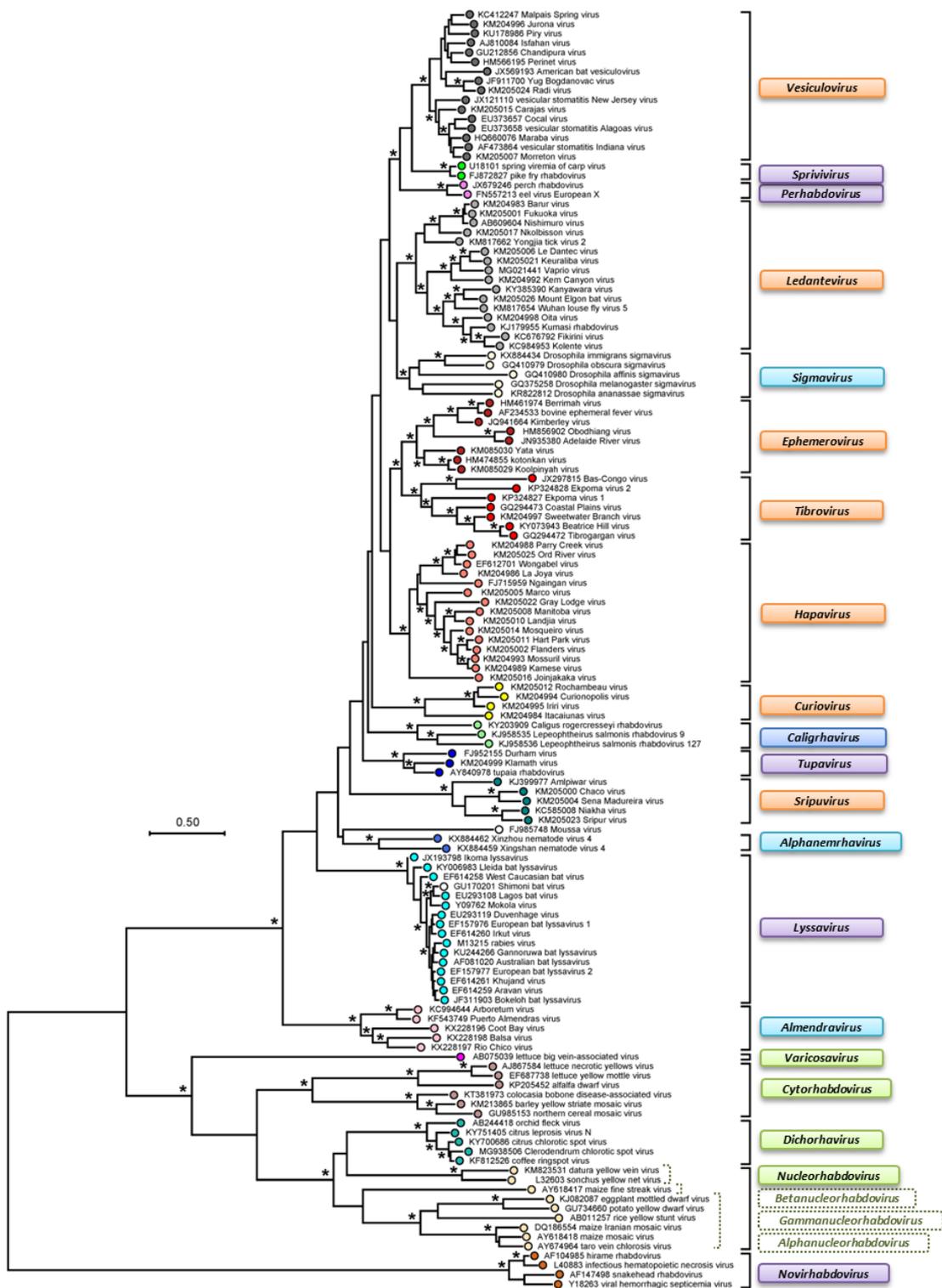
Figure 1.2 Maximum likelihood phylogenetic tree of the family *Rhabdoviridae*. Tree constructed from the full-length L gene sequences of 134 rhabdoviruses assigned to the 20 genera, and one unassigned Rhabdovirus (Moussa virus). Remaining rhabdoviruses were not included as the full-length L gene sequence was not available. Coloured circles at each node differentiate taxa belonging to the different genera, and virus names are preceded by their accession number. Genera names are coloured as per host type; vertebrate hosts are coloured purple, vertebrate hosts with an arthropod vector are coloured orange, invertebrate hosts are coloured blue, and plant hosts are coloured green. The geneus *Nucleorhabdovirus* has recently been split into three genera, indicated by dashed lines and dark green (Dietzgen et al. 2020). The asterisks (*) indicate bootstrap values ≥75 % of 1,000 bootstraps. The scale denotes the number of substitutions per site (0.5; Modified from Walker et al., 2018).

1.2.1.1 Virion Structure

The rhabdovirus virions are enveloped and generally bullet-shaped or bacilliform. Sizes range from 100 - 460 nm in length and 45 - 100 nm in diameter. Classically, cone-shaped viruses infect vertebrates, while plant rhabdoviruses are bacilliform or bullet-shaped. The bullet-shaped LNYV virions are shown in Figure 1.3 (Redinbaugh and Hogenhout 2005; Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018).
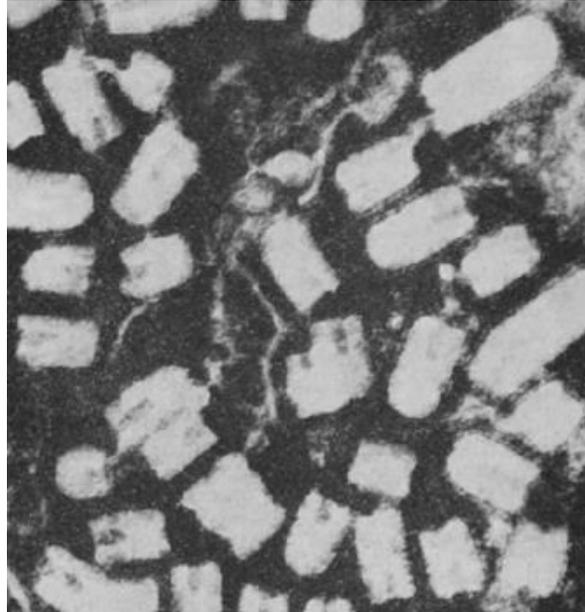
Figure 1.3 Electron micrograph of bullet-shaped LNYV particles. Image magnification is at approximately 91, 500 (Reproduced from Fry et al. 1973).

Glycoprotein peplomer projections, 5 - 10 nm long and ~ 3 nm in diameter, cover the outer surface of the virion, creating a honeycomb pattern on some viruses (Figure 1.4A). The peplomers consist of ectodomain trimers. The ectodomain trimers are considered to be the functional unit for both virion assembly and virus entry into vertebrate and invertebrate host cells (Kuzmin et al. 2009; Walker et al. 2018).

The internal virion structure consists of a helical ribonucleoprotein (RNP) complex. This nucleocapsid structure is 30 - 70 nm in diameter, and cross-striations can be seen when the virion is negatively-stained and viewed under electron micrograph (Dietzgen et al. 2007; Walker et al. 2018). The RNP complex is composed of tightly linked genomic RNA, nucleoprotein (N), RNA-dependent RNA polymerase (L) and polymerase-associated phosphoprotein (P). The matrix protein (M) condenses the RNP complex, encases the nucleocapsid and interacts with the glycoprotein (G) envelope (Figure 1.4B; Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018).
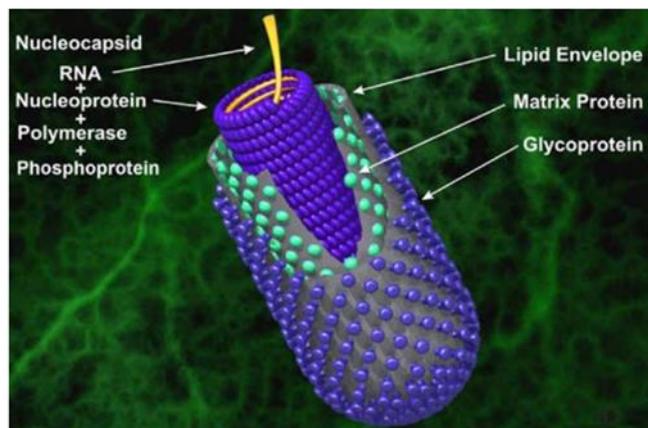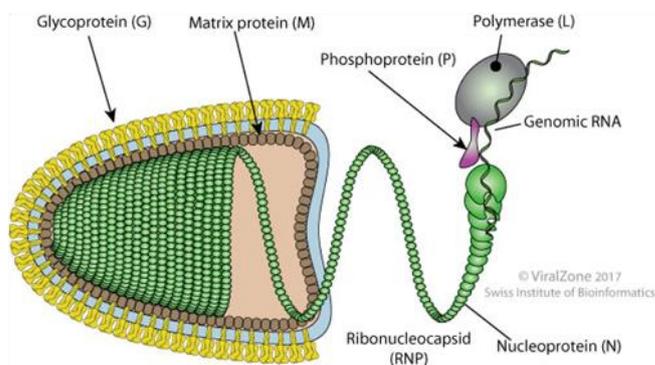
Figure 1.4 Illustrations of rhabdovirus virion and RNP structure. A) Shows virion outer surface with glycoprotein peplomer projections honeycomb pattern, lipid layer imbedded with the matrix protein, and internal helical RNP structure (Reproduced from Dietzgen et al. 2017). B) Shows a cross section of the virion with the RNP structure depicted as unravelling to show the association of the L and P proteins within this helical structure (Reproduced from Hulo et al. 2011; Walker et al. 2018).

1.2.1.2 Genome Structure

The rhabdovirus genome consists of a linear, single-stranded, negative-sense RNA with five canonical genes (from 3' to 5'); N – P – M – G – L. Genomes range from 10.8 kb to 16.1 kb in size, are generally unsegmented, and represent 1 – 3 % of the virion weight (Dietzgen et al. 2017; Walker et al. 2018). The genera *Varicosavirus*, and more recently *Dichoravirus*, have bi-segmented genomes where each RNA segment is encapsulated independently. These have been classified as rhabdoviruses due to substantial sequence identity with known rhabdovirus genera, and their rod shaped morphology (Dietzgen et al. 2017; Walker et al. 2018).

The five canonical genes are framed by regulatory 3' leader and 5' trailer sequences (3' *leader* – N – P – M – G – L – *trailer* 5'), and intergenic regions that include signal sequences for

transcription and translation. Rhabdovirus genomes can also contain further accessory genes (from 5 to 10, or more), and the five canonical genes can also be overprinted or overlapped (Walker et al. 2011; Dietzgen et al. 2017; Walker et al. 2018).

1.2.1.3   Genome Replication and Transcription

The virion RNP complex is active during both transcription and translation. Except for the genera previously known as *Nucleorhabdovirus* and *Dichorhabdovirus*, virus replication occurs in the cytoplasm and follows a common pathway. Host cell entry, un-coating of the endosome, viral transcription and translation, viral genome replication and encapsidation, and lastly, release of the viral genome by budding (Dietzgen et al. 2017). Host cell entry by vector-mediated penetration occurs for most rhabdovirus genera, where the arthropod vector penetrates host cellular walls, and transmits the virus. For those genera that transmit without an arthropod vector (cited in section 1.2.1) host cell entry by receptor-mediated endocytosis is facilitated by the G protein attaching to surface receptors of the cell. Changes in the endosome pH causes the endosome to un-coat, releasing the RNP complex into the cytoplasm (Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018).

After cell entry via either method, the M protein disconnects from the nucleocapsid and initiates viral transcription (Dietzgen et al. 2017). Viral transcription at this stage is also referred to as primary transcription. The RNP complex is used as the viral template, and the RNA genome is repetitively transcribed by the virion transcriptase (Dietzgen et al. 2007; Kuzmin et al. 2009; Walker et al. 2011; Walker et al. 2018). Each gene is repetitively transcribed in a progressive, sequential manner, from 3' to 5' end on a decreasing molar gradient. The transcribed mRNAs are monocistronic, 5'-capped and 3'-polyadenylated (Figure 1.5; Walker et al. 2011; Dietzgen et al. 2017; Walker et al. 2018). The 5' cap is generated by the transcription initiation (TI) sequence which also initiates methylation of the downstream mRNA. The 3'-polyadenylation is generated by the transcription termination polyadenylation (TTP) sequence at the end of each gene, and terminates the mRNA (Figure 1.5; Walker et al. 2011; Dietzgen et al. 2017). The 3' leader sequence lacks the 5' cap, the 3'-polyadenylated (poly A) tail, and is transcribed but not translated (Walker et al. 2018).
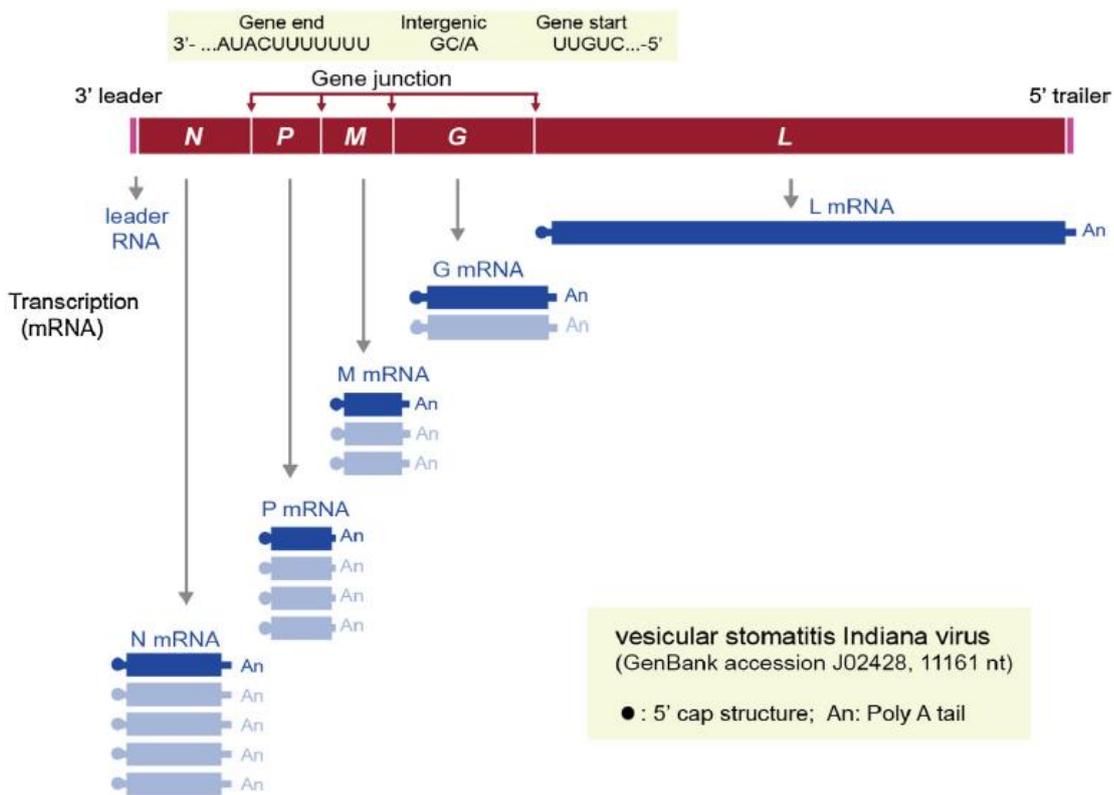
Figure 1.5 Illustration of primary transcription of an example rhabdoviral genome. Vesicular stomatitis Indiana virus is used as an example showing transcription of the five canonical genes, nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and polymerase protein (L). The light blue mRNA depictions represent the decreasing molar gradient of transcribed genes (N > P > M > G > L). The shaded box above the genome representation shows; the 'gene end' sequence or transcription termination polyadenylation (TTP) sequence that generates the 3'-poly A tail on mRNAs and is represented by 'An' on blue mRNAs in this figure, the 'intergenic' sequence that can vary in length and is non-transcribed between the genes, and the 'gene start' or transcription initiation (TI) penta-nucleotide sequence that initiates the 5' cap on mRNAs and is represented by a '•' on blue mRNAs in this figure (Reproduced from Dietzgen et al. 2017).

Using the host cell organelles in the cytoplasm, gene translation occurs from the capped and polyadenylated mRNAs (Kuzmin et al. 2009; Walker et al. 2018). G proteins are transported to the cytoplasmic membrane, and the other viral proteins are expressed by free polyribosomes in the cytosol (Kuzmin et al. 2009). The accumulation of the viral proteins triggers a 'switch' from transcription to genome replication in the polymerase function (Dietzgen et al. 2017; Walker et al. 2018).

As the viral RNA genome is negative-sense, genome replication requires an encapsulated, full-length positive-sense RNA genome to be synthesised. This is referred to as the anti-genome. The anti-genome is initiated by the complementary sequence of the leader RNA and assists the RNP (N/L/P proteins) in encapsidation of the emerging positive-sense RNA that is synthesised by the polymerase complex. The anti-genome binds with the N/P protein complex to function

as a template for synthesis of a full-length, encapsidated, negative-sense RNA genome for viral progeny (Figure 1.6; Dietzgen et al. 2017; Walker et al. 2018).
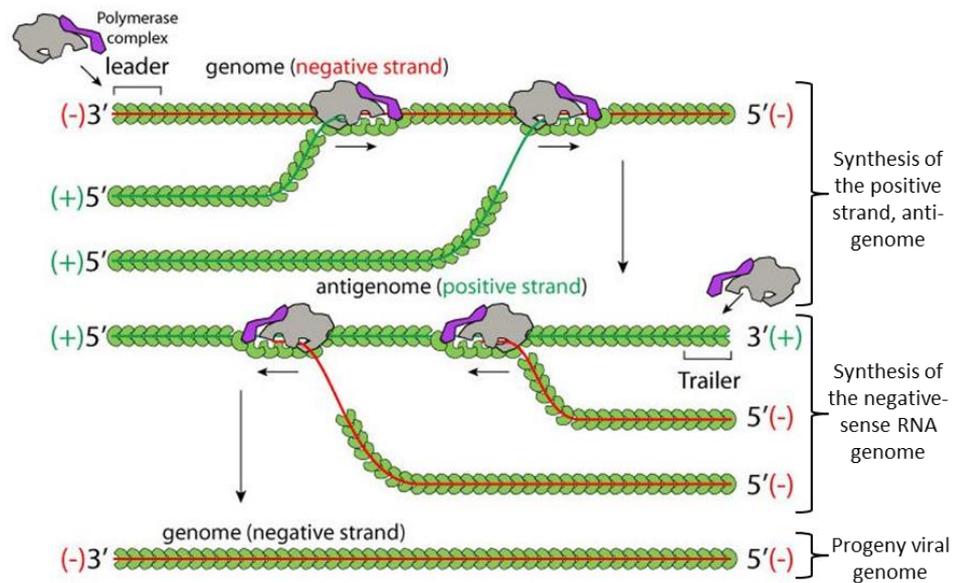
Figure 1.6 Illustration of negative-sense RNA genome replication. Illustration shows the synthesis of the positive strand anti-genome (green strand with blue line) and the negative-sense (olive stand with red line) progeny viral genome. Using the polymerase complex, first the encapsidated full-length anti-genome is synthesised. Using this anti-genome, a full-length, encapsidated, negative-sense progeny RNA genome is synthesised. In this figure the RNP is labelled as 'Polymerase complex' (Modified from Hulo et al. 2011).

Following genome replication, secondary transcription, translation, and replication occurs from the full-length, negative-sense RNA progeny genomes (Mann and Dietzgen 2014; Walker et al. 2018). From these genomes, progeny virions are assembled in a staggered process. The RNP is assembled via M, and lipid envelopes containing G form the virions. The genes are accumulated from different parts of the host cell as the RNP moves to the site of viral budding (Dietzgen et al. 2017; Walker et al. 2018).

The site of viral budding varies depending on the virus and host cell (Redinbaugh and Hogenhout 2005; Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018). For cytorhabdoviruses for example, the RNP accumulates at the endoplasmic reticulum and establishes cytoplasmic thread-like viroplasms. The viroplasms then move to the cytoplasmic membrane sites where G also accumulates. From here, the viroplasms bud into the cytoplasm and form enveloped virions (Redinbaugh and Hogenhout 2005; Mann and Dietzgen 2014; Dietzgen et al. 2017; Walker et al. 2018). Immunocytochemical mapping and analysis of protein interactions for the cytorhabdovirus LNYV have supported this replication model (Martin et al. 2012).

Within the Rhabdovirus family there are six genera of plant rhabdoviruses (Figure 1.2). *Cytorhabdovirus, Dichorhabdovirus, Varicosavirus* (Walker et al. 2018), and those previously known as *Nucleorhabdovirus; Alphanuceorhabdovirus, Betanucleorhabdovirus* and *Gammanucleorhabdovirus* (Dietzgen et al. 2020). The virus strongly correlated to recent lettuce crop losses in NZ, *Lettuce necrotic yellows virus* belongs to the *Cytorhabdovirus* genus.

1.2.2 *Cytorhabdovirus*

The *Cytorhabdovirus* genus was established based on the cytoplasmic site of viral transcription and translation. The classification of this genus has been further supported by phylogenetic analysis (Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007; Kuzmin et al. 2009; Walker et al. 2018). Species classification within the genus is determined by several factors. A nucleotide sequence difference of at least 50 % is required, the plant host range and vector specificity, as well as the ability to be clearly distinguished in serological or by nucleic acid diagnostic tests also influence the classification (Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018). There are currently 11 (Walker et al. 2018) recognised member species, of which there are only seven complete genomes available. There are also 15 unclassified cytorhabdoviruses, of which 11 have their complete genome available (National Center for Biotechnology Information [NCBI] 2004; Walker et al. 2018).

Cytorhabdovirus genomes are unsegmented and range from 12.8 kb – 14.5 kb, encoding six to ten genes (Walker et al. 2018). These include the five canonical rhabdovirus proteins, and an accessory viral movement protein (MP) gene between the P and M genes (Walker et al. 2011; Walker et al. 2018). Some cytorhabdoviruses have been found to encode small additional proteins that have structural similarities to the class 1a viroporins that are detected in mammalian rhabdoviruses (Walker et al. 2018), while LNYV and alfalfa dwarf virus (ADV) have RNA silencing suppressor activity within their P protein. The functions of the other accessory genes found in cytorhabdoviruses are as yet unknown (Redinbaugh and Hogenhout 2005; Walker et al. 2011; Dietzgen et al. 2017; Walker et al. 2018; Whitfield et al. 2018). Figure 1.7 shows the genome organisation, including these accessory genes, of four cytorhabdoviruses.
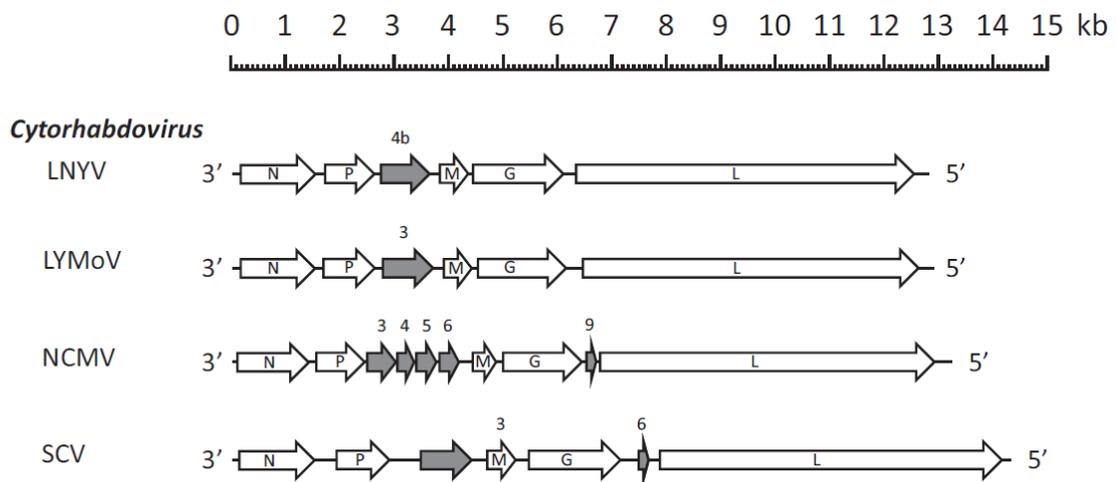
Figure 1.7 Genome organisation of example cytorhabdoviruses. Scale at the top represents genome size in kilobases (kb). Canonical rhabdoviral genes are white, accessory genes are shaded in grey, and the positions of the open reading frames (ORFs) are indicated by the arrow shapes. Cytorhabdoviruses represented include lettuce necrotic yellows virus (LNYV), lettuce yellow mottle virus (LYMoV), northern cereal mosaic virus (NCMV), and strawberry crinkle virus (SCV; Reproduced from Walker et al. 2011).

Overall, cytorhabdoviruses can infect a wide variety of monocotyledon and dicotyledon plant species, with transmission occurring via arthropod vectors. This is primarily *Aphidiae* (aphids), *Cicadellidae* (leafhoppers), and *Delphacidae* (planthoppers; Redinbaugh and Hogenhout 2005; Kuzmin et al. 2009; Walker et al. 2018). In most cases, the virus replicates in the arthropod vector as well as in the host plant. Some cytorhabdovirus species can be transmitted during vegetative propagation, and some can be mechanically transmitted by infected sap. Transmission via seed or pollen has not been reported (Fry et al. 1973; Boakye and Randles 1974; Redinbaugh and Hogenhout 2005; Mann and Dietzgen 2014; Walker et al. 2018).

Lettuce necrotic yellows virus is the type species for the *Cytorhabdovirus* genus. The type species is used as the taxonomic reference for classification (Walker et al. 2018; EzBioCloud 2018). Other discovered cytorhabdoviruses, such as northern cereal mosaic virus (NCMV; Tanno et al. 2000), lettuce yellow mottle virus (LYMoV; Heim et al. 2008), persimmon virus A (PeVA; Ito et al. 2013), ADV (Bejerman et al. 2015), and colocasia bobone disease-associated virus (CBDaV; (Higgins et al. 2016a), were all compared to LNYV to assist with their classification as *Cytorhabdovirus* members (Walker et al. 2018).

1.2.3 Lettuce necrotic yellows virus (LNYV)

1.2.3.1 Infection and Symptoms

LNYV was first officially reported and named in 1963. Suspected lettuce crop losses from the previously unnamed virus have been documented in Australia since 1954 (Stubbs and Grogan 1963). The virus has also been identified in NZ, with suspected lettuce crop losses since 1965, and was first officially reported in 1973 (Fry et al. 1973).

LNYV is transmitted to plants by aphids. Predominately transmission occurs by the blackcurrant-sowthistle aphid *Hyperomyzus lactucae* (Stubbs and Grogan 1963; Fry et al. 1973; Dietzgen et al. 2007)*. The Asian sowthistle aphid *Hyperomyzus carduellins* (Theob.) has also proven to be a vector (Fry et al. 1973; Dietzgen et al. 2007) and, more recently, the blackcurrant-lettuce aphid *Nasonovia ribisnigri* is suspected of transmitting the virus (Stufkens et al. 2002; Walker et al. 2003; Fletcher et al. 2019). LNYV is transmitted in a persistent, circulative and propagative manner, causing the infected aphid to be able to transmit the virus for the rest of its lifetime (Redinbaugh and Hogenhout 2005; Walker et al. 2018; Whitfield et al. 2018).

Virus transmission to the aphid vector occurs via an uninfected aphid feeding on an infected plant and the virus is ingested. In the case of LNYV, it is suspected the uninfected aphid fed on infected sowthistle (*Sonchus* spp.). Virus infection begins in the insect gut, where the viral G peplomers bind to surface receptors on gut epithelial cells, and endocytosis occurs. Here the virus is propagated by viral replication and then circulates through gut basal laminae and muscle cells into nervous tissues, tracheae and insect hemocoels. The insect salivary glands are most likely infected via the hemocoels, but could also be infected via muscle cells or neurons. Virions move into the insect's saliva and are transmitted into other plants. In the case of LNYV this is usually lettuce (*Lactuca sativa;* Carneiro et al. 2005; Redinbaugh and Hogenhout 2005; Gallet et al. 2018; Whitfield et al. 2018). This pathway through the insect vector is depicted in Figure 1.8.

LNYV can also be transmitted transovarially. Studies with the aphid *H. lactucae*, showed transmission across at least two generations (Boakye and Randles 1974; Dietzgen et al. 2007). However, only about 20 % of viruliferous aphids were able to transmit the virus, and only 20 % of these are able to transmit to the next generation (Boakye and Randles 1974; Dietzgen et al. 2007).
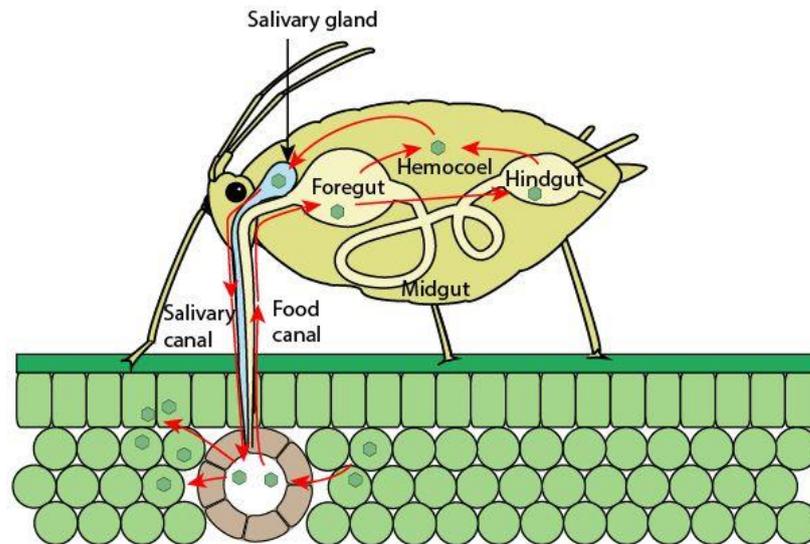
Figure 1.8 Illustration of arthropod transmission of viruses. Circulative propagation of the virus is indicated by the red arrows, showing the feeding arthropod ingesting the virus. The virus infects the insect via the gut, propagates, and circulates to the salivary gland and canal to transmit via the insect's saliva (Carneiro et al. 2005; Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007a; Whitfield et al. 2018; Figure reproduced from Hulo et al. 2011).

LNYV has been shown to infect a variety of plants. This includes *Allium sativum* (garlic), *Arachis hypogaea* (peanut), *Carthamus tinctorius* (safflower), *Cicer arietinum* (chickpea), *Lactuca sativa* (lettuce), *L. serriola* (prickly lettuce or milk thistle), *Lupinus albus* (white lupine), *L. angustifolius* (narrow-leaf lupin), *Medicago polymorpha* (Burclover), and the sowthistle species, *Sonchus oleraceus, S. asper, S. kirkii, S. megalocarpa* and false sowthistle, *Reichardia tingitana* (Plantwise [Date unknown]; Stubbs and Grogan 1963; Fry et al. 1973; Dietzgen et al. 2007a; Higgins et al. 2016b; Fletcher et al. 2019). Sowthistle (*S. oleraceus*) is noted as the most important reservoir host plant and lettuce (*L. sativa*) is the most economically important host (Dietzgen et al. 2007a; Higgins et al. 2016b). For experimental purposes LNYV has also been successfully mechanically inoculated into several plants. These include *Nicotiana glutinosa* and *N. clevelandii, Petunia hybrid* (petunia)*, Datura stramonium* (jimson weed or thorn apple)*, Spinacia oleraceus* (spinach), *Lycopersicon esculentum* (tomato) and *Gomphrena globosa* (globe amaranth; Dietzgen et al. 2007a; Higgins et al. 2016b).

LNYV symptoms on lettuce can vary depending on the age of the lettuce when infected. Lettuce plants infected early become severely stunted (Figure 1.9C). Early infection prior to heart formation causes necrosis of the innermost leaves, often extending to bronzing of the outer leaves following the leaf veins. These symptoms result in the rejection of the plant for sale (Fry et al. 1973; Dietzgen et al. 2007). Later infection causes outer green leaves to appear dull and faded, developing bronzing and necrosis along the veins. Older leaves then become

chlorotic or mottled, and grow in a flattened and flaccid state (Figure 1.9D), commonly progressing to plant death (Stubbs and Grogan 1963; Fry et al. 1973; Dietzgen et al. 2007).

Figure 1.9 Healthy and LNYV infected field lettuce showing a range of symptoms. A) Healthy lettuce. B) Early infection showing stunted growth. C) Early infection showing yellowing of leaves and stunted growth. D) Late infection showing chlorotic outer leaves (Photos provided by Dr Colleen Higgins, Auckland University of Technology).

Other LNYV hosts show a variety of symptoms. Slight stunting and interveinal chlorosis symptoms have been recorded for lupins infected with LNYV, and lethal wilting and necrotic tip burn in infected chickpea. Infected safflower develops mosaic symptoms, and burclover develops chlorotic streaking and slight leaf distortion. LNYV infections of garlic and peanut have been observed in mixed infections with other plant viruses (Dietzgen et al. 2007; Higgins et al. 2016b). In contrast, *Sonchus* spp. appear symptomless, even when they have been successfully mechanically inoculated (Stubbs and Grogan 1963; Fry et al. 1973; Dietzgen et al. 2007).

In mechanically inoculated *N. glutinosa,* symptoms differ across different isolates of LNYV. Systemic leaf crinkling, downcurling, chlorosis or mottling, and vein clearing have been observed (Stubbs and Grogan 1963; Fry et al. 1973; Dietzgen et al. 2007). Mechanical inoculation of lettuce (*L. sativa*) and *Sonchus* spp. has proven difficult; however, *N. glutinosa* is more readily susceptible (Stubbs and Grogan 1963; Crowley 1967; Fry et al. 1973; Dietzgen et al. 2007). Mechanical inoculation techniques aim to replicate the probing of aphid stylets (Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007). This is because when probing for suitable host plants the aphid stylets wound the epidermal plant cells, allowing virus to be directly injected (Boakye and Randles 1974; Redinbaugh and Hogenhout 2005).

It is suspected that lettuce plants are infected by aphids probing. The primary LNYV aphid vector, *H. lactucae,* does not breed or colonise lettuce, so infection is due to infected aphids probing lettuce plants while searching for suitable breeding plants. The aphid overwinters on blackcurrants, then migrates looking for sowthistle species (Fletcher et al. 2019). Indeed, increased incidence of LNYV infection in lettuce plants correlates to aphid migration in late spring and summer (Stubbs and Grogan 1963; Fry et al. 1973; Boakye and Randles 1974; Dietzgen et al. 2007; Diaz et al. 2012; Fletcher et al. 2019).

A survey of NZ aphid populations, from November 2016 through to March 2017, at two lettuce farms, found *H. lactucae* only during November (Fletcher et al. 2019). However, the *N. ribisngri* aphid was found from December through to February, with peak incidence in January. This matches previous records of aphid surveys (Walker et al. 2003; Fletcher et al. 2019). The *N. ribisngri* aphid also overwinters on blackcurrant or gooseberry but in contrast to *H. lactucae,* it emerges to colonise lettuce crops or *Crepis capillaris* (hawksbeard) and *Cichorum intybus* (chicory; Stufkens et al. 2002; Fletcher et al. 2019). *N. ribisngri* colonises lettuce heads and innermost leaves, unlike most aphids that infect lettuce, and arrives early in the growing season (Diaz et al. 2012). *N. ribisngri* is known to cause substantial economic losses in lettuce crops worldwide, and the aphid was discovered in NZ in 2002 (Stufkens et al. 2002; Diaz et al. 2012; Fletcher et al. 2019). A recent project was instigated to confirm *N. ribisngri* as a vector for LNYV in NZ, but unfortunately the original aphid colony collapsed, and another could not be re-established at the time. This was likely due to the warm spring and then the alternating hot and wet summer (Fletcher et al. 2018).

1.2.3.2 LNYV Genome

The LNYV genome contains six genes. The five canonical genes consistent with rhabdoviruses and the MP gene known as 4b positioned between the P and M genes. This is consistent with other cytorhabdoviruses (Dietzgen et al. 2007; Walker et al. 2018). With a genome size of 12,807 nucleotides (nt; Wetzel et al. 1994a), LNYV sits on the smaller end of the size range for current sequenced cytorhabdoviruses (Walker et al. 2018). This genome has an average rhabdovirus coding capacity of 90.4 % (Dietzgen et al. 2007). The six protein coding genes are framed by untranslated, highly complementary 3' leader sequence of 84 nt, and 5' trailer sequence of 187 nt (Dietzgen et al. 2007). These complementary sequences form a "panhandle" structure, as seen in other rhabdoviruses, although the LNYV leader sequence has a two nucleotide overhang (Wetzel et al. 1994a; Ito et al. 2013; Bejerman et al. 2015). The function of the panhandle structure is unclear, but thought to be involved with site recognition signals for the polymerase complex during genome replication (Wetzel et al. 1994a; Bejerman et al. 2015).

The order of genes in the LNYV genome is 3' leader – N – P – 4b (MP) – M – G – L – 5' trailer (Figure 1.10A). The genes are each separated by highly conserved, short, intergenic regions containing the sequence GNU(C/U)(N)nACU where the (N)n represents a variable number of nucleotides (Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007; Walker et al. 2018). Analysis of the viral mRNAs showed their 5' ends are capped with $m^7G^{5'}ppp^{5'}$ (7-methylguanosine) connected via a 5' hydroxyl group by the triphosphate group to the 5' hydroxyl group of the first nucleotide to be encoded (Biology Online 2005). This is consistent with other rhabdoviruses (Dietzgen et al. 2007). The mRNAs encoding the N, M and G proteins also contain additional G and A nucleotides between the 5' cap and the start codon (Wetzel et al. 1994a; Dietzgen et al. 2007a).

After the 3' leader sequence of 84 nt, and an intergenic region of 6 nt, is the start of the N gene. This canonical gene is 1,530 nt in length, with the major ORF 1,374 nt in length. The ORF is flanked by untranslated regions (UTR) of ~ 78 nt, and encodes a protein 459 amino acids long (Figure 1.10A; Wetzel et al. 1994b; Dietzgen et al. 2007a). Three smaller overlapping ORFs have been identified, one in the plus-sense direction (1,227 nt to 1,448 nt) and two in the minus-sense (1,111 nt to 935 and 817 nt to 587 nt), encoding proteins of 57, 59 and 77 amino acids respectively. However, the function of these short length proteins is currently unknown (Wetzel et al. 1994b). Although there are a total of 11 complete cytorhabdovirus genomes available (Walker et al. 2018), there is limited direct RNA sequence similarity across plant rhabdoviruses (Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007).

The N gene is the first gene to be transcribed. As a result, it's mRNA is the most numerous LNYV mRNA in infected cells, the accumulation of which helps trigger the switch from transcription to genome replication (Section 1.2.1.3, Figure 1.5; Callaghan and Dietzgen 2005; Dietzgen et al. 2007a; Walker et al. 2018). This gene, and its translated protein, play an essential role in forming the RNP complex and the infectious nucleocapsid core structure (Section 1.2.1.1). Similar to the RNA genomic sequences, there is limited protein sequence similarity across rhabdoviruses, although the protein folding and RNA binding cavity is highly conserved (Wetzel et al. 1994b; Kuzmin et al. 2009). LNYV and LYMoV have shown the greatest protein similarity, with 56 % similarity of the amino acid sequence that makes up their N proteins (Heim et al. 2008).

The N and P proteins also interact beyond the RNP complex. Prior to the N protein interacting with the viral genome, the P protein acts as a chaperone, helping the N protein retain its soluble form and prevents the N protein from automatically assembling outside of replication (Kuzmin et al. 2009; Martin et al. 2012). Immunocytochemical analysis detecting protein and protein interactions and localisations, has shown these interactions between the N and P proteins (Martin et al. 2012). The N and P interactions have been visualised outside the nucleus and, except for ADV, this was the case for all protein localisation studies of rhabdoviruses (Martin et al. 2012; Bejerman et al. 2015).

For the LNYV N protein, 21 phosphorylation sites have been identified. These sites were shown to be conserved across the LNYV isolates analysed, suggesting functional significance for at least some of these sites (Dietzgen et al. 2007). During a study of several LNYV isolates, characteristics in the sequences of the N gene revealed the isolates could be grouped into two distinct subgroups (Section 1.2.3.3; Callaghan and Dietzgen 2005). Sequence characteristics like the stop codon, the 3' and 5' UTRs, and the intergenic region between the N and P genes supported the groupings, as did phylogenetic analysis (Callaghan and Dietzgen 2005). The phylogenetic analysis and isolate groupings are discussed further in Section 1.2.3.3. An overview of the basic features of all six LNYV genes are illustrated in Figure 1.10B.

A

91 - 1623   2720 - 3765   4412 - 6247

l — | N | P | 4b | M | G | L | — t

1631 - 2712   3773 – 4403   6278 – 12613

91   169   1548   1623   1631   2712

l — | UTR | N | UTR | | P |
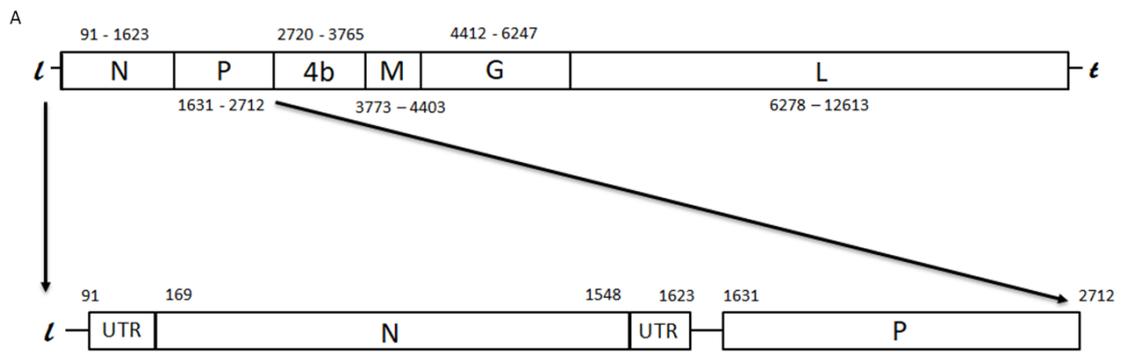
B

Gene	Gene length (nt)	ORF (amino acids)	GenBank accession number	Function	Reference
N	1530	459	L30103	Canonical gene, forms the RNP complex and the nucleocapsid core, interacts with the genomic RNA, and L and P proteins	(Kuzmin et al. 2009; Dietzgen et al. 2017; Walker et al. 2018).
P	1085	300	AF209035	Canonical gene, interacts and chaperones the N protein, interacts with the RNP complex and the L protein as a co-factor, and has shown RNA silencing suppressor activity	(Kuzmin et al. 2009; Mann et al. 2015; Walker et al. 2018).
4b	1046	302	AF209034	Accessory gene, putative viral movement protein, sequence and protein structure similarities with the '30K' superfamily of plant virus movement proteins	(Mann et al. 2015; Dietzgen et al. 2017; Walker et al. 2018).
M	631	177	AF209033	Canonical gene, involved in condensing of the nucleocapsid and forming the bullet-shaped virion, gene expression and the shift between transcription and replication, and interacts with the 4b protein	(Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007a).
G	1836	551	AJ251533	Canonical gene, peplomer spikes bind to surface receptors in arthropod vectors, type 1 glyocprotein involved in signalling the L protein to target the endoplasmic reticulum for cleavage and post-translation modification	(Carneiro et al. 2005; Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007a; Whitfield et al. 2018).
L	6332	2067	AJ746199	Canonical gene, contains characteristic structural motifs common for RNA-dependent RNA polymerases (RdRp), is involved with the RNP complex, and viral RNA replication and transcription	(Redinbaugh and Hogenhout 2005; Dietzgen et al. 2007a; Whitfield et al. 2018).

Figure 1.10 Diagram illustrating the LNYV genome. A) Shows the genome organisation and nucleotide position of the six genes, including the leader (*l*) and trailer (*t*). The N and P gene has been enlarged to show the N gene untranslated regions (UTR) and the black horizontal line representing the intergenic regions between the leader (*l*) sequence and the N and P genes (Figure created by author). B) Shows an overview of the basic features of the six LNYV genes, including the accession numbers for the reference genome and brief points about the gene and protein function for each.

1.2.3.3   LNYV Subgroups and Phylogenetic Analysis

As briefly mentioned earlier (Sections 1.2.1 and 1.2.3.2), phylogenetics has been used extensively to view and aid the understanding of inter-species and intra-species relationships

(Lam et al. 2010). This bioinformatics tool allows the taxonomic and evolutionary relationships to be characterised and evaluated. The idea of comparing molecular sequences to infer relationships was put forward in 1958 by Francis Crick, and then more formally by Emil Zuckerkandl and Linus Pauling in 1965 (Pace et al. 2012). After this, discussions began to focus on the right methodology to infer these relationships and find the right tree (Bininda-Emonds 2009).

Modern methods for phylogenetic analysis can be grouped into two broad categories; distance and discrete methods. Briefly, distance methods calculate the pairwise differences between the sequences involved and progressively builds a tree from the two most closely related sequences. The most common distance method is Neighbour-Joining (NJ), and is considered the fastest method (Baldauf 2003; Bininda-Emonds 2009; Lam et al. 2010). Discrete methods, also known as tree searching methods, act on each sequence directly, looking at each column of an alignment of sequences and searching for the best tree that fits the data. There are several discrete methods, the most popular being the Maximum Likelihood (ML) method (Baldauf 2003; Bininda-Emonds 2009; Lam et al. 2010; Pagán 2018).

The ML method requires the most computer power. This is because it essentially tests every tree possibility, producing a final tree with the greatest likelihood of expressing the data, and that is mathematically preferred over the other tree topologies tested (Baldauf 2003; Bininda-Emonds 2009; Lam et al. 2010; Pagán 2018). This method requires a pre-determined model of evolution – a substitution model suitable to the particular dataset – that provides guidelines regarding base composition, rate, frequency, and nature of substitutions at different sites (Bininda-Emonds 2009; Lam et al. 2010). There are different evolution models available, and the most suitable model can be found prior to tree construction using a model test (Lam et al. 2010).

Both phylogenetic systems can employ the use of bootstrapping to test the robustness, and therefore the confidence, of the final tree. Bootstrapping involves generating random replicates of the sequence alignment and building trees from them. Replicate trees that produce branching patterns that are consistent with the final tree indicate confidence in the final tree topology. Generally, 1,000 bootstraps are performed and values of 50 % or higher are reported, as this indicates at least half of the bootstrap replicates produced the same branch topology (Baldauf 2003; Lam et al. 2010).

Phylogenetics has been used to better understand several aspects of plant viruses, including extensive use in plant virus classification and taxonomy (Sections 1.2.1 and 1.2.2; Pagán 2018;

Walker et al. 2018). It has also provided further insight into the mechanisms that drive evolution, the genetic diversity of plant virus populations, and has informed phylodynamics. Phylodynamics includes the understanding of host defences, the virus mode of transmission, epidemiological dynamics and their interaction with the genetic variation of the plant virus (Pagán 2018).

Phylogenetic analysis of LNYV L gene has supported its classification as a cytorhabdovirus (Dietzgen et al. 2007). While phylogenetic analysis of the LNYV N gene ORF from several isolates formed two monophyletic clades, representing two distinct subgroups for the isolates (Callaghan and Dietzgen 2005). These groupings were named subgroup I and subgroup II (Callaghan and Dietzgen 2005). Further analysis of the N gene ORF at both the nucleotide (Figure 1.11A) and amino acid (Figure 1.11B) levels suggest the LNYV population in both NZ and Australia are made up of two subgroups (Callaghan and Dietzgen 2005; Higgins et al. 2016b).

LNYV is not the only plant rhabdovirus to form subgroups. Genetic variants within viral species has been reported for several plant viruses across the different rhabdovirus plant genera. Phylogenetic analysis of the N and/or L genes have revealed groupings or genetic variants for the alphanucleorhabdoviruses (previously nucleorhabdoviruses; Dietzgen et al. 2020) Taro vein chlorosis virus (TaVCV; Revill et al. 2005) and Eggplant mottled dwarf virus (EMDV; Pappi et al. 2016), and the dichorhaviruses Coffee ringspot virus (CoRSV; Ramalho et al. 2016) and Orchid fleck virus (OFV; Kubo et al. 2009). As well as the cytorhabdoviruses ADV (Samarfard et al. 2018), Raspberry vein chlorosis virus (RVCV; Jones et al. 2019; Rajamäki et al. 2019), SCV (Klerks et al. 2004; Koloniuk et al. 2018), and the suspected cytorhabdovirus Ivy latent virus (IvLV; Petrzik 2012).

Variation studies and phylogenetic analysis of the LNYV N gene has displayed distinct topographic characteristics between the two subgroups. Isolates in the subgroup I clade showed longer branches with high bootstrap values, while isolates in the subgroup II clade had shorter branches with bootstrap values below 50 % and therefore were considered not well supported (Figure 1.11A; Higgins et al. 2016b; Ajithkumar 2018).

Sequence variation was detected across the isolates. Including within each subgroup, and not just between the two subgroups, at both the nucleotide and amino acid levels (Callaghan and Dietzgen 2005; Higgins et al. 2016b; Ajithkumar 2018). Higgins et al. (2016b) found the subgroup I isolates were 93.5 % - 100 % identical at the nucleotide level, and pairwise difference values were calculated to be 0.0 – 0.228 by Ajithkumar (2018). Isolates belonging to

subgroup II were 97.2 % - 99.6 % identical at the nucleotide level (Higgins et al. 2016b), and pairwise difference values were between 0.002 – 0.218 (Ajithkumar 2018). Less variation was observed at the amino acid level, with subgroup I showing 98.3 % - 100 % identical (Higgins et al. 2016b), with pairwise difference value of 0.0 – 0.019 (Ajithkumar 2018), and subgroup II showing 98.8 % - 100 % identical (Higgins et al. 2016b), with pairwise difference value of 0.0 – 0.01 (Ajithkumar 2018).

Phylogenetic trees constructed with protein sequences also reflected less variation, compared with nucleotide trees. Less branches were observed for both subgroup clades (Figure 1.11B). This apparent conservation of the N protein sequences may be due to a need to conserve the structure of the protein. Specific functions of the N protein may require precise protein structures to enable these functions to be carried out (Callaghan and Dietzgen 2005; Kuzmin et al. 2009). In fact, most of the observed nucleotide changes have resulted in synonymous substitutions (Callaghan and Dietzgen 2005), conserving the protein sequence. This has also been observed in other plant rhabdoviruses, such as ADV (Samarfard et al. 2018), CoRSV (Ramalho et al. 2016), OFV (Kubo et al. 2009) and EMDV (Pappi et al. 2016).

Phylogenetic analysis which included LNYV N gene isolates from NZ and Australia identified noteworthy relationships with the isolates. Within both subgroups the NZ and Australian isolates formed country specific subclades (Figure 1.11A, and Figure 1.11B; Higgins et al. 2016b; Ajithkumar 2018). In addition, the Australian isolates revealed an unexpected population structure. Of the 18 Australian isolates collected after 1993, none belonged to subgroup I, suggesting this subgroup has become extinct in Australia (Callaghan and Dietzgen 2005; Higgins et al. 2016b). Meanwhile in NZ, both subgroups have been identified in both the North and South Islands (Higgins et al. 2016b; Ajithkumar 2018; Fletcher et al. 2019). In fact, both subgroups have been isolated from the same host species, from the same farm, in the same year (Higgins et al. 2016b; Ajithkumar 2018). This was also seen with the last Australian subgroup I isolate in 1993 (Callaghan and Dietzgen 2005), and suggests the two subgroups have co-existed and evolved independently across space and time (Callaghan and Dietzgen 2005; Higgins et al. 2016b).

The origin and timing of the emergence of LNYV is currently unclear. Bayesian Evolutionary Analysis Sampling Trees (BEAST; https://beast.community) analysis suggested that subgroup I emerged earlier (150 years ago) than subgroup II (75 years ago; Higgins et al. 2016b). These estimates have been supported by the earlier emergence of subgroup I observed in nucleotide tree analysis (Figure 1.11A; Higgins et al. 2016b; Ajithkumar 2018). Conversely, phylogenetic trees from LNYV N gene protein sequences (Figure 1.11B), and whole genome analysis with

other rhabdoviruses (Ajithkumar 2018), indicated the subgroup II lineage emerged earlier than the subgroup I lineage (Higgins et al. 2016b; Ajithkumar 2018). However, this branching pattern does not reflect the current LNYV subgroup population dynamic in Australia, where subgroup I appears to have died out. If the subgroup I population emerged after subgroup II, why would it then have died out early in the 1990's?

The current hypothesis is supported by N gene nucleotide trees and the BEAST analysis. Higgins et al. (2016b) suggests that subgroup II emerged more recently and rapidly, dispersing quickly and supplanting subgroup I. Subgroup II was possibly able to do this as the result of more-optimal interactions with insect and/or plant hosts (Higgins et al. 2016b). Meanwhile in NZ, the continued country-wide presence of both subgroups suggests subgroup II has yet to displace subgroup I (Higgins et al. 2016b). Although with recent increased lettuce crop losses in NZ, it has been suggested that subgroup II could be exhibiting a stronger presence, or a more severe strain of LNYV may have arrived in NZ from Australia (Fletcher et al. 2019). Sampling of lettuce crops in the areas experiencing these losses may provide insight into the source, add support for hypothesises regarding these losses, and update the current LNYV subgroup population structure in NZ. Additional samples would provide more LNYV sequences for analysis and may further elucidate the relationships between the LNYV subgroups.
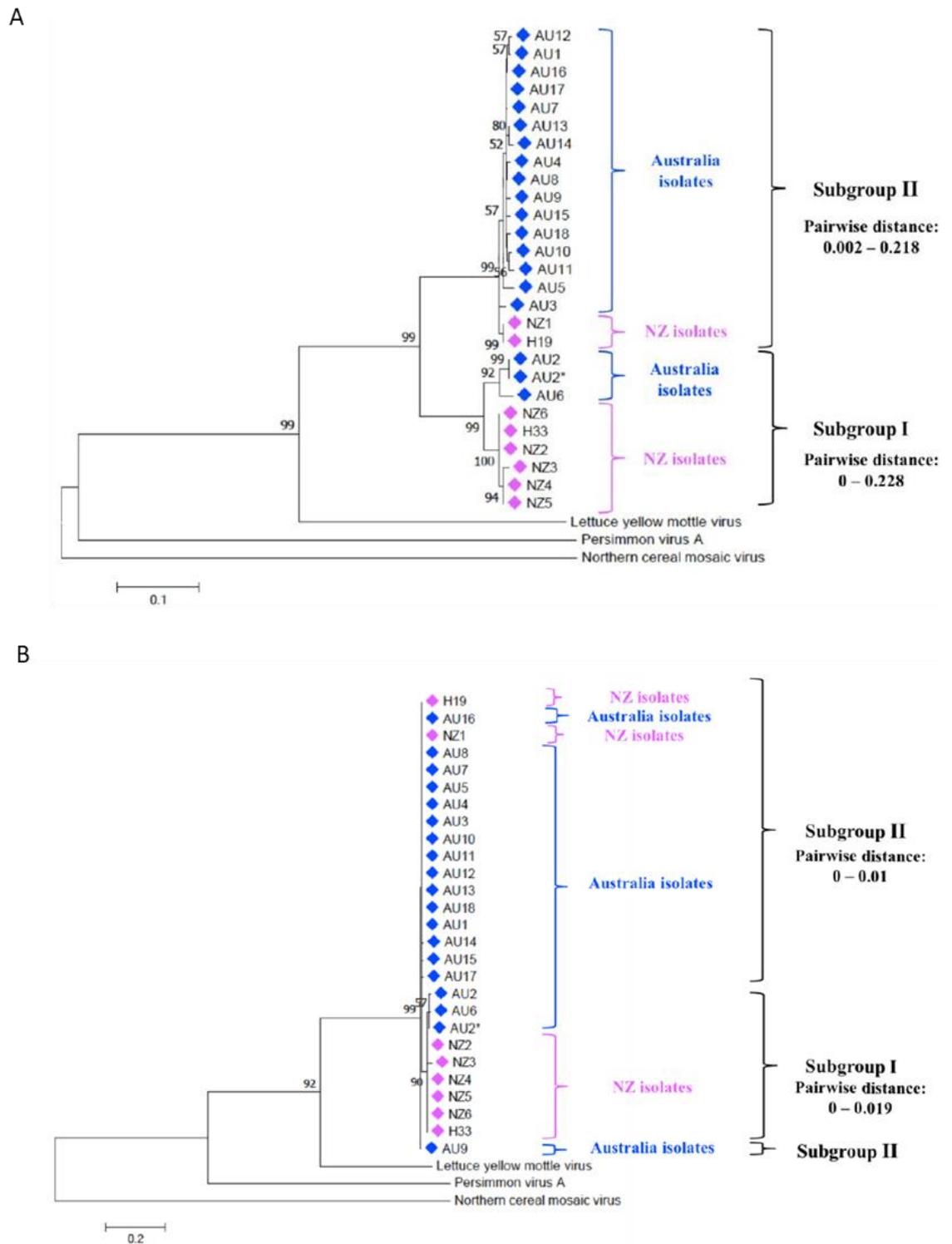
Figure 1.11 Maximum-likelihood (ML) phylogenetic trees of the LNYV N gene. A) Shows ML phylogenetic analysis of LNYV N gene nucleotide sequences. B) Shows ML phylogenetic analysis of LNYV N gene protein sequences. Isolate country of origin is indicated with NZ isolates coloured pink and Australian isolates coloured blue. Subgroups indicated in black with pairwise distance values shown. N gene sequences from three cytorhabdoviruses were included as outgroup sequences and are named in full. Scales represent the number of substitutions per site (Reproduced from Ajithkumar 2018).

1.2.3.4 LNYV Mixed Infections and Quasi-species Model

NZ plant pathogen records document ten viruses that infect lettuce (Veerakone et al. 2015), and pathogen surveys of lettuce crops often identify plants infected with multiple viruses (Fletcher et al. 2005; Fletcher et al. 2019). Mixed infections have also been detected in other plants (Schneider and Roossinck 2001; Jo et al. 2017; Koloniuk et al. 2018). Mixed lettuce infections in NZ can include the viruses *Beet western yellows virus* of the genus *Polerovirus*, family *Luteoviridae*, *Cucumber mosaic virus* of the genus *Cucumovirus*, family *Bromoviridae*, *Lettuce mosaic virus* and *Turnip mosaic virus* of the genus *Potyvirus*, family *Potyviridae*, *Mirafiori lettuce big-vein virus* of the genus *Ophiovirus*, family *Aspiviridae*, and *Lettuce big-vein associated virus* and LNYV of the genera *Varicosavirus* and *Cytorhabdovirus* of the family *Rhabdoviridae* (Fletcher et al. 2005; Veerakone et al. 2015; Fletcher et al. 2019).

In 2017 two types of mixed infections were each detected in approximately 20 % of symptomatic lettuce heads (Fletcher et al. 2019). Mixed infections consisted of LNYV and Cucumber mosaic virus, and LNYV and Lettuce big-vein disease – caused by *Mirafiori lettuce big-vein virus* and *Lettuce big-vein associated virus* (Fletcher et al. 2019). This shows that mixed infections are generally viruses of different genera, although they can belong to the same virus family, such as the rhabdoviruses lettuce big-vein associated virus and LNYV.

As discussed in Section 1.2.3.3, variant strains of several rhabdoviruses have been identified. However, reports of mixed or co-infected hosts with rhabdovirus variants are limited to a very recent report of OFV (Roy et al. 2020), SCV (Koloniuk et al. 2018), and Viral haemorrhagic septicaemia virus (VHSv) and Infectious hematopoietic necrosis virus (IHNv) in fish hosts (Benmansour et al. 1997; Thompson et al. 2011; Pierce and Stepien 2012). Studies have shown that mixed infections could have a synergistic relationship, influencing disease symptoms, and may play a role in viral evolution, the levels of diversity within the viral population, and viral adaptability (Martinez et al. 2012; Pierce and Stepien 2012; Jo et al. 2017; Leeks et al. 2018). These properties appear to fit a quasi-species model, which is characteristic of RNA viruses (Pierce and Stepien 2012).

The quasi-species model describes the complex and dynamic distribution of viral variants, or mutants, forming a viral population for a virus species. This viral population, also known as a 'swarm' or 'cloud', is made up of viral genome sequences that are closely related, but not identical, to the master, reference, or 'wild type' viral sequence genome (Figure 1.12; Compans et al. 2006; Domingo 2010; Martinez et al. 2012; Viruses are models… 2018). The viral quasi-species model was first demonstrated about 40 years ago, and has provided an important

mathematical framework for understanding the population genetics and structure of rapidly evolving viruses (Compans et al. 2006; Viruses are models… 2018).
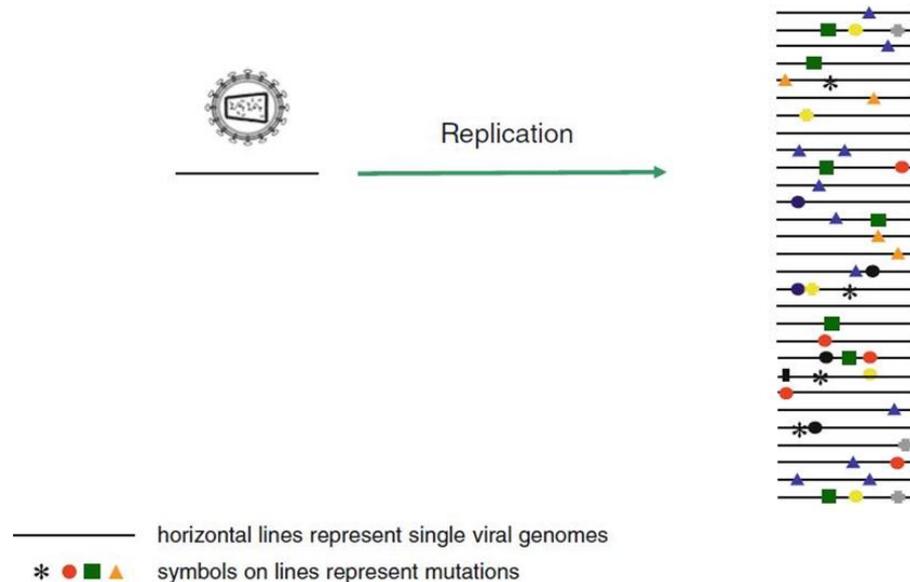
Figure 1.12 Representation of viral quasi-species model. Illustration depicting a single viral genome, represented by the black horizontal line, producing viral variants displaying mutations, represented by the coloured symbols, after replication (Reproduced from Martinez et al. 2012).

The quasi-species model is characteristic of RNA viruses due to their small genome size and high mutation rates that, therefore, generate high genetic diversity. High mutation rates allow RNA viruses to evade the antiviral defences of hosts and adapt quickly to new environments (Eigen 1993; Martinez et al. 2012; Pierce and Stepien 2012). The RNA-dependent RNA polymerase protein of RNA viruses, known as the L protein in rhabdoviruses, lacks the ability to efficiently 'proof-read' during replication (Compans et al. 2006; Domingo 2010; Martinez et al. 2012). This causes high, approximately 10^{-4}, nucleotide substitutions per site per replication cycle (Martinez et al. 2012; Pierce and Stepien 2012). Mutations can also occur via RNA recombination and reassortment in the presence of mixed infections in the same host cell (Domingo 2010; Martinez et al. 2012). Reassortment is limited to segmented viral genomes, and has been shown to contribute to the genetic diversity of the dichorhavirus, OFV (Kondo et al. 2017; Roy et al. 2020). Recombination can occur in all viruses, although, it reportedly occurs consistently less frequently in negative sense single-stranded RNA viruses (Domingo 2010; Díaz-Martínez et al. 2018). However, it has been suggested as a possible mechanism behind rhabdoviruses gaining accessory genes (Walker et al. 2011).

Within the viral quasi-species population, the variant sequences are subjected to continuous genetic variation by selective pressures and viral fitness (Figure 1.13; Compans et al. 2006; Domingo 2010; Martinez et al. 2012). Viral fitness is regulated by the viral phenotypes that determine survival and the production rate of viable viral progeny (Eigen 1993; Compans et al. 2006; Domingo 2010). Viral fitness is further defined into different facets; replicative fitness, which refers to the capacity of the virus to produce viable viral progeny in the given environment, transmission fitness, which refers to the capacity of the virus to transmit to new hosts for survival due to the host's defences and finite lifespan, and epidemiologic fitness, which refers to the capacity of the virus to become the dominant variant, phenotype or clade, in respect to the other variants, phenotypes or clades of the same virus in the given environment (Wargo and Kurath 2012). Viral variants with advantageous traits become the dominant variant, phenotype or clade through positive selective pressures leading to a new consensus genome sequence and an increase in fitness (Figure 1.13). Negative selective pressures or bottleneck events cause an increase in traits that modify the consensus genome sequence in an attempt to increase viral fitness. However, this can actually lead to decreased fitness, and the elimination or decrease in frequency of a variant genotype (Figure 1.13; Compans et al. 2006; Domingo et al. 2012; Martinez et al. 2012). Plant viruses, in particular, undergo frequent bottleneck events as they overcome plant and insect host defences and barriers to transmission (Kuzmin et al. 2009; Domingo et al. 2012; Dietzgen et al. 2016).
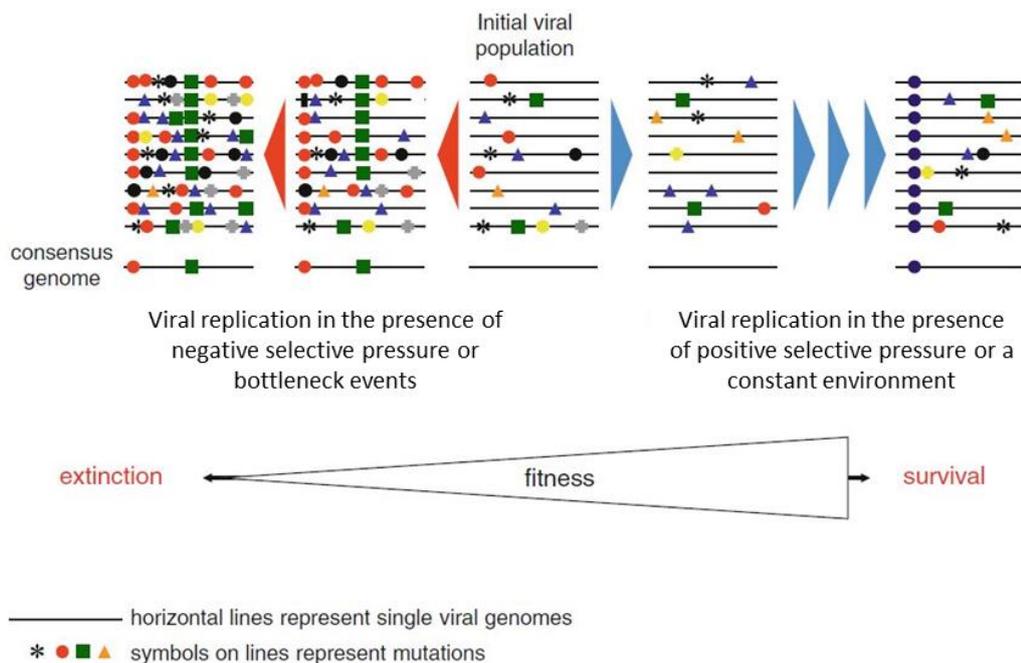
Figure 1.13 Representation of quasi-species viral fitness and survival. Illustration depicting the effects of positive selective pressure, blue arrows on the right, and negative selected pressure, red arrows on the left, on an initial viral population. Unrestricted replication in a constant environment after multiple replication cycles, multiple blue arrows, leads to fitness gain and survival, depicted by the triangle at the bottom. Negative selected pressure or bottleneck events leads to decreased fitness and the accumulation of mutations that modify the consensus sequence, depicted by the symbols on the consensus genome. The consensus genome refers to the master, reference or 'wild type' viral sequence for the population (Modified from Martinez et al. 2012).

The quasi-species model has been heavily explored in relation to drug-resistance mutations and mutagenesis for human RNA viruses such as human immunodeficiency virus (HIV), influenza virus, and hepatitis B and C (Domingo et al. 2012; Martinez et al. 2012; Viruses are models… 2018). Several vesiculoviruses that infect vertebrates, have demonstrated the quasi-species model (Jordan et al. 2000; Novella et al. 2013), as have the novirhabdoviruses, VHSv and IHNv, that infect fish (Benmansour et al. 1997; Thompson et al. 2011; Pierce and Stepien 2012). However, these appear to be the only published reports of rhabdovirus species demonstrating the quasi-species model.

The quasi-species model has been demonstrated in plants through studies of host-virus interactions (Schneider and Roossinck 2001; Ali and Roossinck 2017; Jo et al. 2017). A range of hosts and viruses have been studied, which has demonstrated that universal conclusions regarding the population dynamics of RNA quasi-species cannot be made. The RNA quasi-species population dynamics are likely to be different for different viruses and virus strains in different hosts (Ali and Roossinck 2017). This is supported by studies showing that quasi-species population size for the same virus was not constant, and depended on the host

environment, including plant hosts of different cultivars (Schneider and Roossinck 2001; Jo et al. 2017). In fact, it has also been shown that this change in host environment generated different rates of mutation and evolution, as well as the level of population variation and diversity in the same virus (Schneider and Roossinck 2001; Jo et al. 2017). However, a study analysing quasi-species populations in single and mixed infections found that the population variation could be stable and was not particularly affected by the host (Ali and Roossinck 2017).

This study also found that different viral strains generated different quasi-species population variation. Single infections showed higher population variation compared to mixed infections, suggesting that mixed infections may cause shifts in population diversity levels through the modification of viral evolution rates (Ali and Roossinck 2017). It is clear that general conclusions regarding the population dynamics of quasi-species cannot be made at this stage (Ali and Roossinck 2017). Nevertheless, recent technological advancements with high-throughput sequencing (HTS) allows for greater sequence coverage of the quasi-species genome sequences, compared to traditional methods (Jo et al. 2017). This may allow for further studies and greater understanding of the host-virus population dynamics of quasi-species.

The mixed virus infections seen in NZ lettuce are all RNA viruses. Although it has not been confirmed, a quasi-species nature for LNYV is likely to exists as it is an RNA virus. The quasi-species model also suggests mixed infections may well have contributed to the variation seen in LNYV, though there is no evidence of this reported at this stage. LNYV has shown sufficient variation to give rise to the two subgroups, and variability within these subgroups is apparent from phylogenies, particularly for subgroup II (Section 1.2.3.3). The sequence variability for either subgroup within a host has not yet been reported, and there has not been any report describing co-infection of the two subgroups, suggesting that they are discrete populations. Although this is an important consideration for understanding the LNYV subgroup population structure, it is beyond the scope of this study.

1.2.4 Virus Detection

As discussed earlier in Section 1.1, plant pathogens have a significant influence on crop losses globally. It is estimated the costs due to infections in the United States alone are US$40 billion annually (Fang and Ramasamy 2015). In NZ, horticulture greatly supports the NZ economy, with 5,500 commercial fruit and vegetable growers employing about 60,000 people (HortNZ

2017). NZ exports 60 % of what it grows, with horticultural exports tripling in the last 20 years to approximately NZ$5.6 billion, almost 10 % of NZ's total export merchandise (Beresford and Mckay 2012; HortNZ 2017; PFR 2019). However, biosecurity issues due to plant pests have significant cultural, social, environmental, and economic impacts (Greer and Saunders 2012; Ganley and Bulman 2016).

A key example of a costly plant pest is grapevine leafroll disease (GLD). GLD is a complex of virus species from the family *Closteroviridae*, and is present in all grape-growing regions of the world, including NZ (Almeida et al. 2013; Veerakone et al. 2015). It is transmitted by mealybug species, *Pseudococcidae*, and soft scales, *Coccidae*. Although specific economic data for GLD is limited, it has been estimated to range from US$25,000 to US$40,000 per vineyard hectare for those vineyards with a 25-year lifespan (Almeida et al. 2013). NZ currently has 35,000 hectares of vineyard producing over NZ$1.6 billion in exports (PFR 2019). There is no cure for GLD so management of infection levels within a vineyard is the only way for growers to maintain economic sustainability. Infection levels above 1 % means infected vines need to be removed, termed roguing, and replaced with vines tested to be disease free (Almeida et al. 2013; PFR 2017). The cost of roguing to growers per hectare for ultra-premium Chardonnay vines is approximately NZ$160,000 (PFR 2017). This highlights the importance of virus detection methods for querying infection, monitoring infection levels, and screening plants and seeds prior to planting, export and/or import (Gachon et al. 2004; Almeida et al. 2013; Fang and Ramasamy 2015; Fox and Mumford 2017).

The very nature of viruses - their very small size, composition and pathogenesis - requires understanding of their aetiology for effective management (Jeong et al. 2014; Selvarajan and Balasubramanian 2016; Fox and Mumford 2017). Viral infection symptoms including crinkling, mosaic, browning or necrotic areas on leaves, could be due to other causes such as unfavourable weather or nutrient imbalances (Almeida et al. 2013; Jeong et al. 2014; Dobhal et al. 2016). In addition, some viral infections could be asymptomatic, making diagnosis of viral infection by symptoms more difficult than for other plant pathogens, and agrochemical control treatments are not an effective option, particularly after infection (Jeong et al. 2014; Dobhal et al. 2016). Therefore, accurate identification and diagnosis is an important first step for crop management systems (Jeong et al. 2014; Fang and Ramasamy 2015; Selvarajan and Balasubramanian 2016; Fox and Mumford 2017; Pallás et al. 2018).

Plant virology can be traced back to the late 19th Century. Although the development of methods for characterising and differentiating viral diseases didn't start until the 1930's (Fox and Mumford 2017). This led to attempts to catalogue the world's viruses (Fox and Mumford

2017), and the creation of the International Committee on Taxonomy of Viruses (ICTV) during the International Congress of Microbiology held in Moscow in 1966 (ICTV 2019). As of the tenth report from ICTV, an excess of 1,400 plant virus species have been listed (Fox and Mumford 2017; ICTV 2019).

Understanding the aetiology of plant viruses has also influenced the diagnostic detection technologies for these pathogens. The specific characteristics of viruses that mean they cannot be cultured or directly observed has necessitated virologists being more prepared to explore new detection methods (Fox and Mumford 2017). Although this does not mean these technologies are embraced for frontline applications, trends in the methods of detection have progressed alongside the development of new technologies (Figure 1.14; Boonham et al. 2014; Fox and Mumford 2017). This has been demonstrated in a review of first and novel plant virus detections in the United Kingdom (UK) by Fox and Mumford (2017). Within the review period, 1980 to 2014, virus detection technology evolved from those involving biology, morphology, and serology, to those involving molecular, sequencing, and next generation sequencing (NGS) or high-throughput sequencing (HTS) technologies (Figure 1.14; Fox and Mumford 2017).
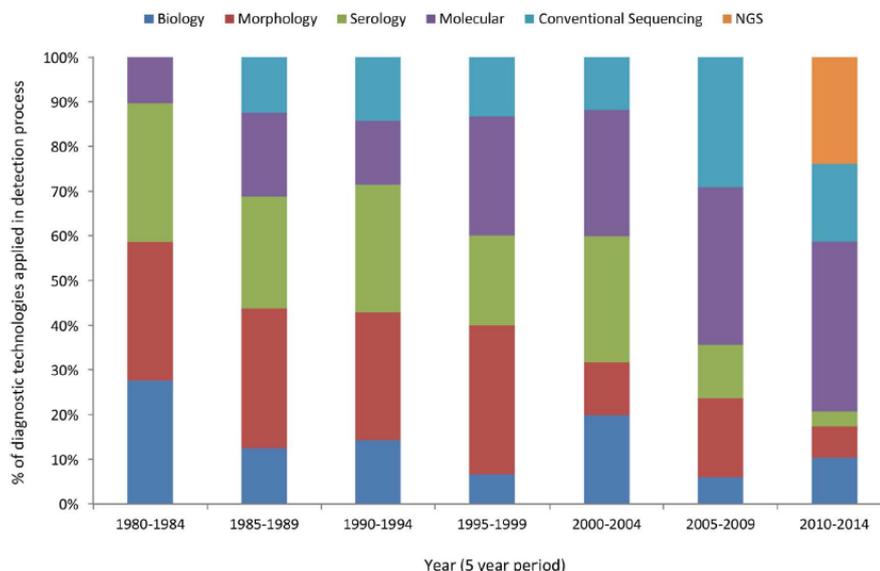
Figure 1.14 Technologies applied in detection and diagnosis of first detection and novel discoveries of viruses, viroids and virus-like agents in the United Kingdom. Data was collected from reports where the use of that technology has led to a positive finding. Figures are the percentage of findings where each type of methods had been reported to have been used in the detection of a new finding. Data presented as periods of five years. NGS refers to next-generation sequencing (Reproduced from Fox and Mumford 2017).

Traditionally, initial diagnosis of plant virus infections involved analysis of their 'biology' or pathological mechanism. This has included analysis of field symptoms and symptoms on mechanically or graft-inoculated indicator hosts (Schaad and Frederick 2002; Veerakone et al. 2015; Selvarajan and Balasubramanian 2016; Fox and Mumford 2017). This is reflected by a contribution of almost 30 % to first virus discoveries between 1980 and 1984 in the UK (Figure 1.14; Fox and Mumford 2017). However, analysis of the pathological mechanism has severe limitations, with viral symptoms on different hosts showing variation or even asymptomatic, and some viruses being unable to be transmitted by mechanical sap inoculation (Dietzgen et al. 2007; Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016).

The further development of electron microscopy (EM) and transmission electron microscopy (TEM) allowed the inclusion of 'morphology' in diagnostic reports. This technology dominated detection techniques in the 1980's and 1990's (Figure 1.14; Fox and Mumford 2017). However, these methods required specialist knowledge and years of experience (Boonham et al. 2014).

From the 1990's there is a distinct up-take in the use of specific targeted technologies such as serology (Section 1.2.4.1) and molecular methods (Section 1.2.4.2; Fox and Mumford 2017). The report by Fox and Mumford (2017) also noted the use of two or more methods to confirm findings, and suggested that the sudden rise of NGS (or HTS) techniques in 2010 to 2014, appeared to be replacing 'molecular' and 'conventional sequencing' methods in first detection reports (Figure 1.14).

1.2.4.1 Serology Technologies

Serology technologies are target specific as they rely on the specific antibody-antigen binding relationship. Viral antiserum is produced by injecting warm-blooded mammals, such as rabbits or goats, with purified virus or viral antigen over a period of several weeks. The polyclonal immunoglobulins, also known as antiserum, is then purified from the mammal's serum and is used for virus specific detection (Clark and Adams 1977; Jeong et al. 2014; Selvarajan and Balasubramanian 2016). Several technologies utilised this antibody-antigen binding relationship, in particular, the enzyme-linked immunosorbent assay (ELISA) which included an antibody that was conjugated or 'linked' to an enzyme or 'reporter' molecule and produced a measurable product (Clark and Adams 1977; Cox et al. 2012).

The ELISA method became popular after Clark and Adams (1977) developed the method to function in a microplate, allowing for large numbers of samples to be tested. There are several ELISA formats, of which the sandwich or double antibody sandwich ELISA (DAS-ELISA) were the

most popular for plant virus detection (Clark and Adams 1977; Torrance and Jones 1981; Boonham et al. 2014; Fox and Mumford 2017). In the sandwich ELISA method (Figure 1.15), the viral antiserum polyclonal antibodies are attached to a solid surface, the microplate, and are used to 'capture' viral antigens in the sample to be tested. The sample to be tested is added along with a second enzyme or reporter linked antibody and incubated to allow antibody-antigen binding. The second antibody is used for detection and measurement, and binds to different sites or epitopes on the viral antigen compared to the capture antibody. The viral sample is then 'sandwiched' between the two antibodies. Lastly, an enzyme or reporter substrate is added to generate the measurable product (Figure 1.15). Conventionally, a colour change was observed, and the optical density measured. Other reporter types can include a fluorophore and fluorescence is measured or a chemiluminescent substrate is used and the luminescence is measured. The reporter signal measurement is proportional to the amount of viral antigen present in the sample, allowing for a semi-quantitative measurement (Clark and Adams 1977; Torrance and Jones 1981; Cox et al. 2012; Fang and Ramasamy 2015).
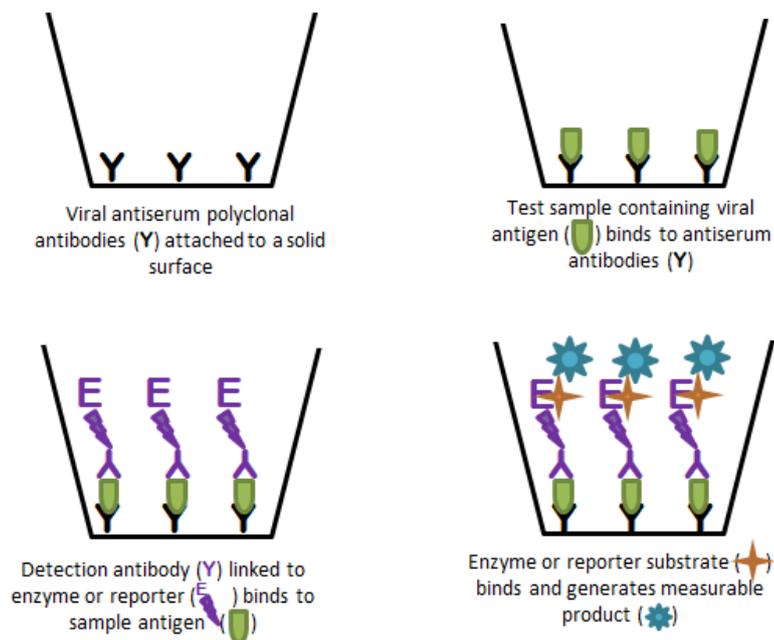
Figure 1.15 Diagram illustrating the major steps involved in sandwich ELISA. Viral antiserum is attached to a solid surface, usually a microplate. Test sample is added and binds to antiserum. Detection antibody with linked enzyme or reporter binds to the test sample antigen, effectively 'sandwiching' the test sample antigen between two antibodies. Enzyme reporter substrate is added and generates a measurable product (Figure created by author based on Clark and Adams 1977).

Target specific serological technologies proved to be much faster than morphology and biological symptom diagnosis methods. This contributed to their high and consistent use, seen between 1980 and 2004 (Figure 1.14; Boonham et al. 2014; Fox and Mumford 2017). The commercialisation of antiserum, equipment and reagents meant serological methods were cost effective, sensitive, simple and versatile (Boonham et al. 2014; Jeong et al. 2014). Only a small amount of antibody is required for testing and the process can be semi-automated, allowing a large number of samples to be tested at once (Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016). The addition of specific monoclonal or recombinant antibodies increases assay performance through improved sensitivity and specificity (Boonham et al. 2014; Fang and Ramasamy 2015). However, the overall process is still lengthy, with specialised equipment, expertise, and a large purified amount of the viral sample is needed to generate the specific antiserum (Boonham et al. 2014; Jeong et al. 2014). Although the antibody-antigen binding relationship is specific, cross-reactivity and false positive results can occur, and the method is often not appropriate for closely related virus strains or when a wide range of pathogens need to be screened for (Boonham et al. 2014; Jeong et al. 2014; Fang and Ramasamy 2015; Selvarajan and Balasubramanian 2016).

Serological technologies to detect LNYV appear to have been widely researched from late 1960's to early 1990's (Dietzgen et al. 2007). This research was inspired in 1967 with the development of a method to purify the virus (McLean and Francki 1967). Although this method produced highly purified LNYV virus particles, it produced only a small amount, while requiring a large amount of infected plant material (McLean and Francki 1967). A further improved purification method was reported in 1989 (Francki et al. 1989). The early serological technologies involving tube precipitin or gel diffusion tests required highly concentrated virus and could produce non-specific precipitates (Francki and Randles 1970; McLean et al. 1971; Francki et al. 1989). The N and G genes were able to be detected using monoclonal antibodies from leaf material by immunoblot detection and indirect ELISA (Dietzgen and Francki 1988; Dietzgen and Francki 1990). However, the development of a DAS-ELISA proved to be significantly more sensitive (Dietzgen et al. 2007), and was able to detect LNYV in leaf material and aphid vectors (Chu and Francki 1982). LNYV diagnosis by ELISA was used in a study assessing the spatial patterns of LNYV and Lettuce big-vein disease in lettuce crops in Australia (Coutts et al. 2004).

Throughout this time, difficulties with extracting the virus due to degradation from reducing agents and solvents was documented (McLean and Francki 1967; Francki and Randles 1970; Francki et al. 1989; Dietzgen and Francki 1990). Although LNYV can be detected by DAS-ELISA, no serological difference between LNYV strains has been detected (Francki et al. 1989), further

supporting the unsuitability of serological technologies to distinguish closely related viruses (Boonham et al. 2014; Jeong et al. 2014).

1.2.4.2 Molecular Technologies

During the same time period when serological technologies were being established, some of the first molecular technologies were also developed. As seen in Figure 1.14, molecular and conventional sequencing technologies were common technologies used for first virus detections from 1985 to 2014 (Fox and Mumford 2017). The Sanger sequencing (Section 1.2.4.2.1) method in 1977 (Sanger et al. 1977), and the polymerase chain reaction (PCR; Section 1.2.4.2.3) in 1986 (Mullis et al. 1986), represent two of the foundation technologies of modern molecular, sequencing, and HTS technologies.

1.2.4.2.1 *Sanger Sequencing*

By the 1960's many proteins had been sequenced. However, this was a laborious process of fragmenting the RNA, separation of fragments by chromatography and electrophoresis, followed by sequential exonuclease digestion of each individual RNA fragment, with the RNA sequence determined from overlaps (Shendure et al. 2017). In 1977 two methods to decode sequences were reported; chemical degradation (Maxam and Gilbert 1977) and chain-termination (Sanger et al. 1977). Of which Sanger's chain-termination method ultimately proved to be the most useful and is now known as Sanger sequencing (Stranneheim and Lundeberg 2012; Shendure et al. 2017).

Briefly, Sanger sequencing involves the use of the enzyme, DNA polymerase. The DNA polymerase is used to extend a complementary nucleotide strand, or chain, using deoxynucleotide triphosphates (dNTPs) initiated by an oligonucleotide primer. The reaction also included dideoxynucleotide triphosphates (ddNTPs) of each nucleotide base that lacked a 3' hydroxyl group, causing the termination of the extension of the complementary strand. The arbitrary incorporation of the chain-terminating ddNTPs meant that complementary stands were of different lengths, eventually generating complementary strands that had terminated at each template nucleotide position. The strands were separated by electrophoresis, enabling each template nucleotide to be sequenced (Sanger et al. 1977; Stranneheim and Lundeberg 2012; Shendure et al. 2017; Bluth and Bluth 2018). Modern Sanger sequencing is automated, with the template strands differentiated by capillary gel electrophoresis, and the ddNTPs are

labelled by distinct fluorescent dyes. The fluorescent signals are then read by a laser and specific software creates a chromatograph or electropherogram (Figure 1.16; Stranneheim and Lundeberg 2012; Shendure et al. 2017; Bluth and Bluth 2018).
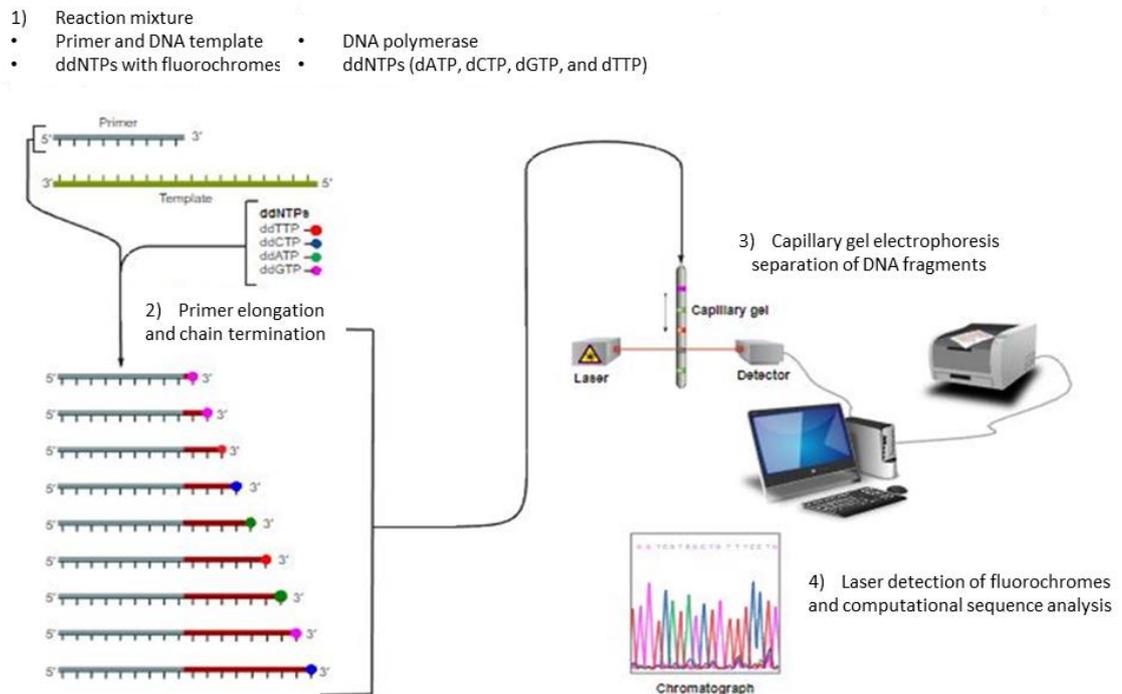
Figure 1.16 Illustration of modern Sanger sequencing. 1) Shows reaction components. 2) Shows the different lengths of the complementary strands and the incorporation of fluorescently labelled ddNTPs. 3) Shows the separation of the complementary strands by capillary gel electrophoresis. 4) Shows fluorescently labelled ddNTPs read by laser and computational visualisation by chromatograph or electropherogram (Modified from Bluth and Bluth 2013).

Although Sanger sequencing came into immediate use, its role in virus detection was generally characterising and confirming viruses (Fox and Mumford 2017). After initial identification by symptoms and negative ELISA tests of known viruses, new viruses were cloned and sequenced to identify their genome organisation (Boonham et al. 2014). This sequence information was extremely useful for other molecular technologies, such as PCR (Section 1.2.4.2.3), and more recently isothermal amplification (Section 1.2.4.2.6; Jeong et al. 2014; Selvarajan and Balasubramanian 2016; Fox and Mumford 2017). As seen in Figure 1.14, conventional sequencing contributed to about 10 % of first virus reports from 1985 to 2004, with an increase to 30 % for the period 2005 to 2009 (Fox and Mumford 2017). The Human Genome Project, completed in 2004, incentivised many improvements to Sanger sequencing technologies during the 1990's, accumulating in the modern methodology illustrated in Figure

1.16 (Stranneheim and Lundeberg 2012; Shendure et al. 2017). At the same time alternative sequencing techniques were also explored, leading to the development of HTS and NGS technologies (Section 1.2.4.2.2) arising not long after the Human Genome Project was completed (Stranneheim and Lundeberg 2012; Selvarajan and Balasubramanian 2016; Shendure et al. 2017).

For many years Sanger sequencing was considered the gold standard for sequencing. This is in most part due to its high accuracy, long sequence reads, and cost-effectiveness in regards to labour, instruments and reagents (Stranneheim and Lundeberg 2012; Shendure et al. 2017; De Palma 2018 Nov). Sanger sequencing is ideal for sequencing specific genes or loci (De Palma 2018 Nov). This sequencing data can then be used for further bioinformatic analyses such as phylogenetic analysis (Callaghan and Dietzgen 2005; Bejerman et al. 2015; Dietzgen et al. 2017; Kondo et al. 2017), mutation and strain variant identification (Kondo et al. 2017; Samarfard et al. 2018), and commonly, confirmation or validation of HTS results (Section 1.2.4.2.2; Stranneheim and Lundeberg 2012; Boonham et al. 2014; Khalifa et al. 2015; Pallás et al. 2018; Roy et al. 2018; Samarfard et al. 2020), to list a few.

The Sanger sequencing method was used to determine both the genome organisation (Wetzel et al. 1994a) and the genomic sequence (Wetzel et al. 1994b; Dietzgen et al. 2006) for LNYV. With this information available, molecular technologies, such as PCR (Section 1.2.4.2.3), rather than serological DAS-ELISA methods (Section 1.2.4.1), are able to be used for virus detection (Thomson and Dietzgen 1995). Sanger sequencing information from different LNYV isolates has also been used for phylogenetic analysis and further analyses of the six LNYV genes (Callaghan and Dietzgen 2005; Martinez et al. 2013; Mann et al. 2015; Higgins et al. 2016b; Mann et al. 2016b; Mann et al. 2016a).

1.2.4.2.2 *High Through-put Sequencing (HTS) and Next-Generation Sequencing (NGS) Technologies*

HTS technologies were initially referred to as NGS. However, HTS technologies are now referred to as 'second' or 'third' generation sequencing technologies, due to the evolving nature of technology and the changing title of which technologies are of the 'next generation' (Stranneheim and Lundeberg 2012). HTS technologies have different methodologies compared to Sanger sequencing, although they still utilise the polymerase enzyme, and prior amplification or preparation of the target DNA is still required (Egan et al. 2012; Stranneheim and Lundeberg 2012; Shendure et al. 2017). Multiple different technologies were released

after completion of the Human Genome Project, of which some are no longer utilised (Shendure et al. 2017). The HTS technologies that proved most popular include sequencing by synthesis (SBS) with reversible dye termination used by Illumina, sequencing by ligation used by Complete Genomics, ion-sensitive SBS used by Ion Torrent, and single-molecule sequencing used by Pacific Biosciences and Oxford Nanopore Technologies, which are also referred to as 'third generation' sequencing technologies (Egan et al. 2012; Stranneheim and Lundeberg 2012; Shendure et al. 2017). Although these technologies produce shorter read lengths compared with Sanger sequencing, approximately 200 bp compared to approximately 900 bp, their clonal amplification and template libraries allow for rapid, massively parallel, high-throughput, specific sequencing (Figure 1.17; Egan et al. 2012; Stranneheim and Lundeberg 2012; Shendure et al. 2017).
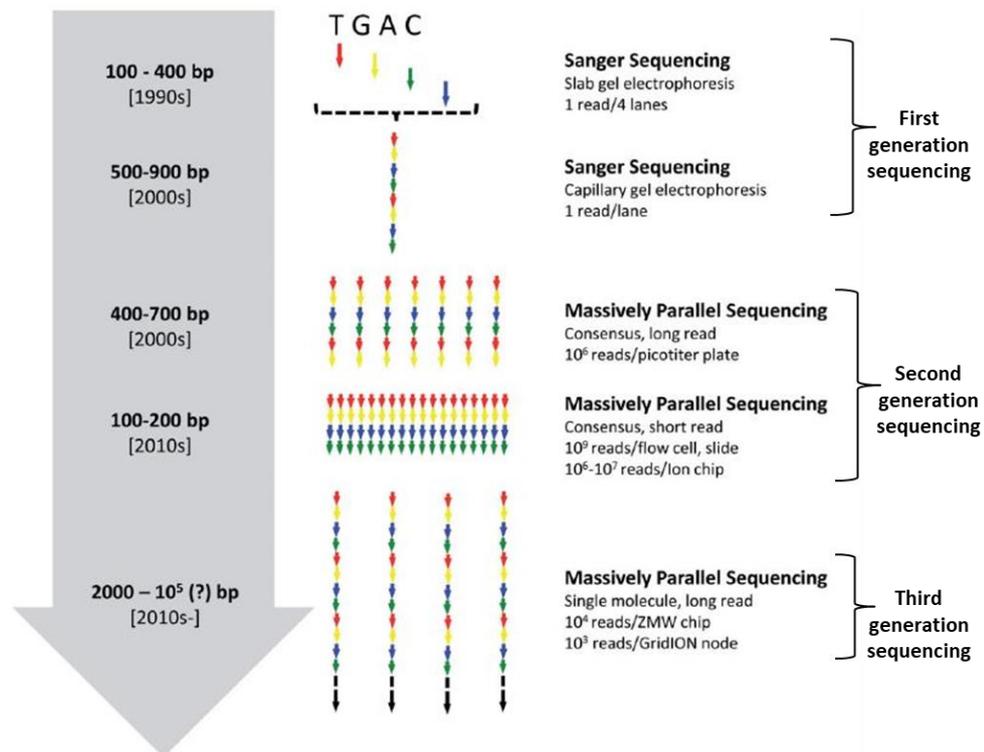
Figure 1.17 Diagram representing the development of molecular sequencing technologies. The key changes in read length and degree of parallelism are illustrated from the 1990's to 2010's. Grey downward arrow on the left shows progression of read length in base pairs (bp) and date in square brackets. Coloured arrows represent the four DNA bases. Under second generation sequencing 'picotiter plate' refers to the pyrosequencing technology no longer utilised, 'flow cell, slide' refers to Illumina sequencing technology, and 'Ion chip' refers to Ion Torrent sequencing technology. Under third generation sequencing 'ZMW chip' refers to Pacific Bioscience sequencing technology, and 'GridION node' refers to the sequencing technology from Oxford Nanopore Technologies (Modified from Stranneheim and Lundeberg 2012).

HTS technologies can provide broad-spectrum, unbiased detection of even closely related genomic sequences, without requiring prior sequence knowledge (Selvarajan and Balasubramanian 2016; Pallás et al. 2018). This makes these technologies ideal for the screening and detection of pathogens, and have been used for plant viral detection studies, with several reports released in 2009 (Boonham et al. 2014; Adams and Fox 2016). As seen in Figure 1.14, between 2010 and 2014, first detection reports in the UK using 'NGS' contributed to almost 30 % of first discoveries of plant viruses (Fox and Mumford 2017). However, the key advantages of NGS or HTS technologies – rapid, massively parallel, and their high-through-put - also generates massive amounts of data.

Initially, complicated bioinformatic tools and expertise were required to decipher and understand this data. Although general bioinformatic pipelines and tools have since been developed and are openly available, they are still computer intensive, require some bioinformatic, rather than biological, expertise, and different pipelines can give differing results depending on the HTS technology and methodology used (Egan et al. 2012; Boonham et al. 2014; Adams and Fox 2016; Selvarajan and Balasubramanian 2016; Pallás et al. 2018). In addition, separation of the viral, or pathogen, sequence data from that belonging to hosts, other organisms present, or any cross-contamination, is required (Adams and Fox 2016; Pallás et al. 2018). Several pre-analysis and sample preparation approaches to balance this have been developed depending on the HTS technology and what information is required from the testing, but each has its own disadvantages (Egan et al. 2012; Boonham et al. 2014; Adams and Fox 2016).

Another drawback when using HTS technologies for broad-spectrum genome sequencing, is the discovery of viruses outside those being screened for (Adams and Fox 2016). The discovery of novel viruses which may or may not cause symptoms is problematic, as well as the discovery and potential characterisation of novel viruses based on nucleotide or amino acid sequence similarity to known viruses, without any supporting biological data (Boonham et al. 2014; Adams and Fox 2016; Pallás et al. 2018). These drawbacks have contributed to the continued 'validation' of HTS results with Sanger sequencing and PCR technologies, and their use in more established labs compared to frontline applications (Stranneheim and Lundeberg 2012; Boonham et al. 2014; Khalifa et al. 2015; Pallás et al. 2018; Roy et al. 2018; Samarfard et al. 2020). Although costs for HTS technologies are decreasing (Boonham et al. 2014; Adams and Fox 2016), the preparation steps, analysis, amount of data generated, and deciphering this data, make the sequencing and detection of known viruses using HTS an expensive option that has yet to be routinely implemented (Boonham et al. 2014; Adams and Fox 2016; Selvarajan and Balasubramanian 2016; Pallás et al. 2018). Viruses where sequence data or conserved

genomic regions are already available enable viral infections to be detected by more cost effective technologies such as PCR (Section 1.2.4.2.3; Boonham et al. 2014; Jeong et al. 2014; Adams and Fox 2016; Selvarajan and Balasubramanian 2016; Pallás et al. 2018).

Despite the drawbacks discussed here, the use of HTS technology has recently detected two variant strains of two different cytorhabdoviruses. Plants suspected of RVCV infection did not amply a diagnostic product when tested with an established PCR based test, so HTS was used (Jones et al. 2019). The HTS reads, with PCR and Sanger sequencing filling in any gaps, identified two variant genomes of RVCV. From this data PCR based testing with degenerate primers for detection of both RVCV strains was developed (Jones et al. 2019). For the other cytorhabdovirus, SCV, the complete genome sequence was sought from symptomatic plants. The sequences produced from HTS identified two variant strains of SCV. Current SCV detection targets the L gene, which was previously the only published SCV sequence. From the HTS genomes the development of new PCR based detection focusing on the other SCV genes can be developed (Koloniuk et al. 2018).

Recently HTS technology was also used to sequence the LNYV genome of two NZ isolates, both a subgroup I and subgroup II isolate (Ajithkumar 2018). Previously only an Australian subgroup I genome was available (Dietzgen et al. 2006). This data was then used for *in silico* analysis of the G gene (Ajithkumar 2018). To date there are no reports of the use of HTS technology for the specific detection of LNYV, however, several other molecular detection methods have been developed and these are discussed in the subsequent Sections (Sections 1.2.4.2.3 and 1.2.4.2.6; Higgins et al. 2016b; Ajithkumar 2018; Fletcher et al. 2018).

1.2.4.2.3 *Polymerase Chain Reaction (PCR)*

Developed in 1986 (Mullis et al. 1986), PCR quickly became an essential element for molecular technologies (Schaad and Frederick 2002; Jeong et al. 2014). PCR was able to be applied to other biological and molecular applications such as cloning, gene expression and manipulation, genotyping and sequencing, and pathogen detection, to name a few (Schaad and Frederick 2002; Jeong et al. 2014). This technology proved to be extremely versatile, and its core function was to amplify a specific sequence of DNA, generating millions of identical copies of the target DNA (Jeong et al. 2014; Bluth and Bluth 2018).

PCR is essentially a chemical reaction, with similar components to Sanger sequencing. The technology also utilises the DNA polymerase enzyme to incorporate dNTPs to extend a specific sequence of the target DNA (Bluth and Bluth 2018). The specific target sequence of the DNA is

identified by a pair of short oligonucleotide primers, representing the 5' forward and 3' reverse sequences (Jeong et al. 2014; Bluth and Bluth 2018). The specific DNA sequence is amplified by 30 to 40 repeated cycles of three steps (Figure 1.18). Step one consists of denaturing the DNA strands at temperatures of 94 °C or above. Step two consists of annealing the oligonucleotide primers at temperatures of 50 – 70 °C, and then step three consists of extension of the primers at ~ 70 °C by the incorporation of the dNTPs via the DNA polymerase (Mullis et al. 1986; Jeong et al. 2014; Bluth and Bluth 2018). As the reaction occurs sequentially, the target sequence, guided by the primers, is amplified exponentially generating millions of identical DNA copies (Jeong et al. 2014; Bluth and Bluth 2018).
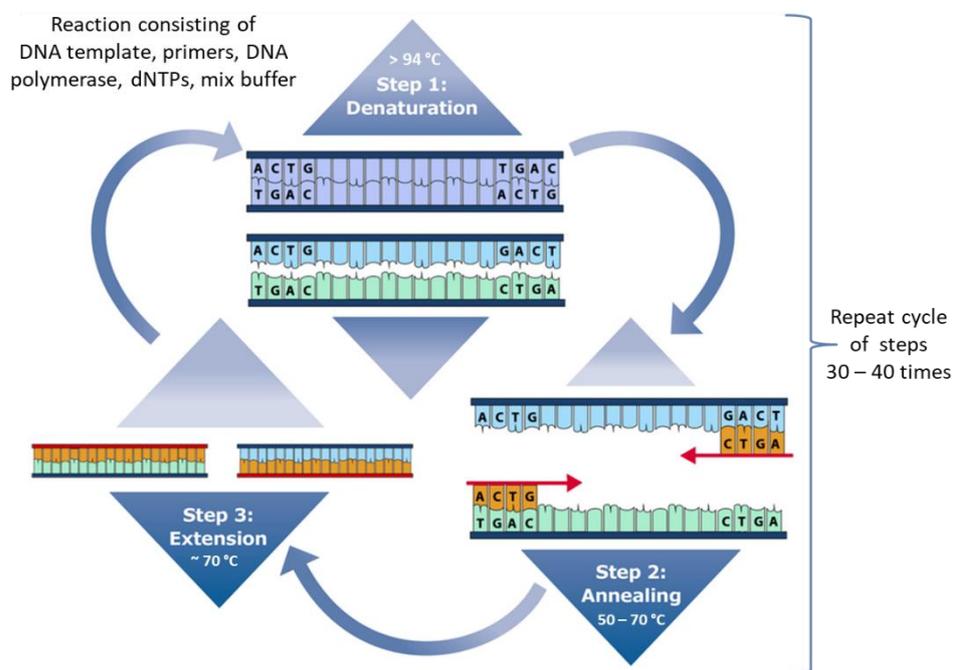
Figure 1.18 Illustration of the polymerase chain reaction (PCR) process. Step 1) Denaturation of the DNA template into two strands at 94 °C or above. Step 2) Annealing of the oligonucleotide primers (coloured orange) to each strand at 50 – 70 °C. Red arrows indicate primer direction on each strand. Step 3) Extension of the primers using deoxyribonucleotide triphosphates (dNTPs; coloured orange) by the DNA polymerase enzyme. This cycle of steps is repeated 30 – 40 times (Modified from University of Waikato 2017)

After the PCR process was completed, correct target amplification was confirmed using size discrimination. This was achieved by agarose gel or capillary gel electrophoresis (Bluth and Bluth 2013; Boonham et al. 2014). The amplified target sequences could then be further utilised in other molecular applications such as cloning or sequencing (Boonham et al. 2014; Jeong et al. 2014).

Technological improvements contributed to increasing accessibility and speed of the PCR process, as well as greatly lowering costs. Such improvements include the development of thermostable *taq* DNA polymerase; from the bacterium *Thermus aquaticus*, programmable and portable thermocycler instruments, and the standardisation of reagents (Lorenz 2012; Boonham et al. 2014; Jeong et al. 2014; Bluth and Bluth 2018). PCR technology also proved to be highly sensitive, specific, cost-effective and flexible (Gachon et al. 2004; Boonham et al. 2014; Jeong et al. 2014; Bluth and Bluth 2018). When compared to serological methods such as ELISA (Section 1.2.4.1), PCR assays were easier and faster to establish, with higher sensitivity and accuracy (Gachon et al. 2004; Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016).

Further development of this technology has increased its versatility for use with other nucleic acids and molecular applications. To amplify and detect multiple nucleic acid targets in the same reaction, multiplex PCR can be performed by incorporating primer pairs for each specific target (Boonham et al. 2014; Jeong et al. 2014; Pallás et al. 2018). Multiplex PCR does require careful primer and assay design using prior knowledge of the target nucleic acid sequence, but allows for highly cost effective, simultaneous, rapid, and potentially broad-spectrum amplification and detection of the target sequences (Boonham et al. 2014; Jeong et al. 2014; Pallás et al. 2018). The addition of a step prior to the regular PCR cycles allows for the amplification of RNA samples using the enzyme, reverse transcriptase. Referred to as reverse transcription PCR (RT-PCR), a complementary DNA (cDNA) strand is synthesised from the RNA template, followed by the regular PCR process (Jeong et al. 2014; Bluth and Bluth 2018). The development of real-time or quantitative PCR (qPCR; Section 1.2.4.2.4) in the 1990's provided the ability to quantify and measure the target nucleic acid sequence as it was being amplified (Higuchi et al. 1993; Gachon et al. 2004; Fox and Mumford 2017; Pallás et al. 2018). Numerous other advances with the PCR process have been achieved. However, those mentioned here are the most relevant to plant virus detection (Boonham et al. 2014; Jeong et al. 2014; Pallás et al. 2018).

Plant virus detection using PCR was first documented in 1989 (Puchta and Sänger 1989; Vunsh et al. 1990). Puchta and Sänger (1989) demonstrated that extremely low amounts of hop stunt viroid (HSVd) could be amplified by RT-PCR. Following this, in 1990 Vunsh et al. demonstrated the first use of RT-PCR to detect an encapsidated virus, *Bean yellow mosaic virus,* from plant tissue and confirming the use of RT-PCR as a diagnostic tool for plant virology. From here PCR became a well-established and popular method for virus detection (Gachon et al. 2004; Boonham et al. 2014; Jeong et al. 2014). As seen in Figure 1.14, molecular technologies

contributed around 10 % from the 1980's, jumping to 30 % in the later half of the 1990's, for first detections of plant viruses (Fox and Mumford 2017).

After the genome mapping and sequencing of the LNYV N gene in 1994, data was available for the development of a molecular detection assay for LNYV (Wetzel et al. 1994a; Wetzel et al. 1994b). Using PCR primers developed by Wetzel et al. (1994b), Thomson and Dietzgen (1995) reported a diagnostic assay for the detection of plant viruses. They demonstrated molecular detection of LNYV from both lettuce and *N. glutinosa* leaves using RT-PCR (Thomson and Dietzgen 1995). The LNYV N gene is the first gene to be transcribed, resulting in the accumulation of N gene mRNA in infected cells (Section 1.2.1.3; Callaghan and Dietzgen 2005; Dietzgen et al. 2017). This therefore, makes it an ideal target sequence for molecular virus detection (Callaghan and Dietzgen 2005; Dietzgen et al. 2007; Walker et al. 2018). Over time, several RT-PCR primer sets able to diagnose LNYV have been developed that focus on the N gene, and these are discussed in greater detail in Chapter 2.

1.2.4.2.4 *Quantitative Polymerase Chain Reaction (qPCR)*

The concept of qPCR was developed to reduce downstream processes and cross-contamination after PCR (Higuchi et al. 1992; Higuchi et al. 1993). In 1992, Higuchi et al. (1992) reported a closed tube PCR system using the addition of ethidium bromide (EtBr) dye to signal the accumulation of dsDNA and therefore, the amplification of the target DNA. This method reduced cross-contamination and eliminated the need for post-PCR gel electrophoresis (Higuchi et al. 1992). In 1993, Higuchi et al. reported the real-time monitoring of the kinetics of PCR using a camera to monitor the fluorescence intensity. Fluorescence increased as the EtBr bound to the exponentially increasing dsDNA in the reaction (Higuchi et al. 1993).

From this, three kinetic phases of PCR were identified. The exponential phase represented the exponential amplification of the target DNA. The linear phase represented the slowing of the reaction as reaction reagents were used up, and the plateau phase represented the end of the reaction, where no more amplification was taking place (Figure 1.19; ThermoFisher [Date unknown]; Higuchi et al. 1993; Bluth and Bluth 2018). It was found the fluorescence intensity was directly related to the amplification of the target DNA, allowing for the target DNA to be quantified and measured in real time (Higuchi et al. 1993).
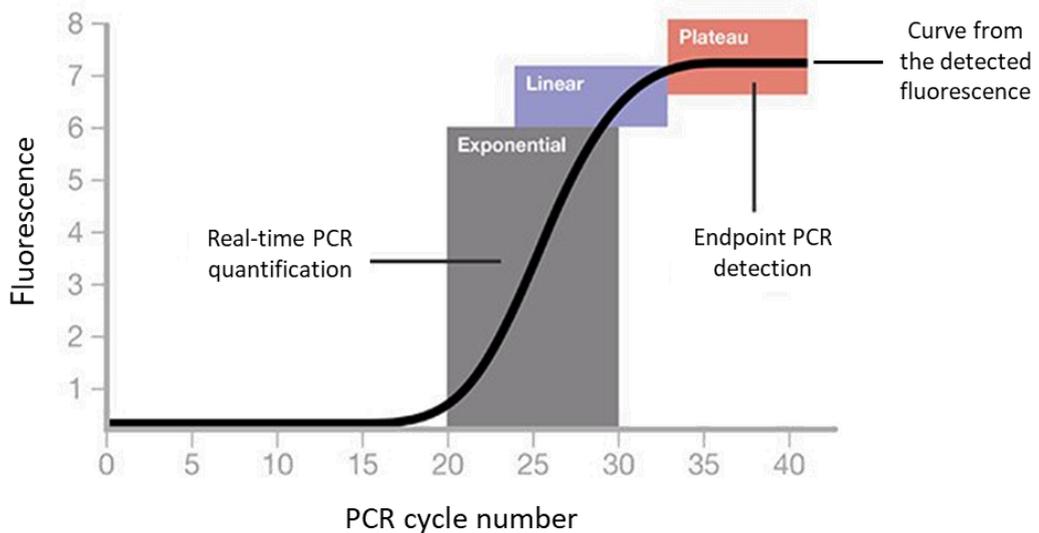
Figure 1.19 Illustration of PCR kinetics seen during qPCR. This illustration represents the amplification curve generated during qPCR. The thick black curve represents the distinctive shape from the detected fluorescence. The PCR kinetics are described in three stages as indicated on the illustration (Modified from ThermoFisher Scientific [Date unknown])

Real-time or quantitative PCR, provided a single, closed-tube system (Higuchi et al. 1993; Applied Biosystems 2009; Jeong et al. 2014). This reduced cross-contamination, post PCR processing, time, labour, and costs (Gachon et al. 2004; Applied Biosystems 2009; Jeong et al. 2014). The assay proved to be simple, rapid, convenient, and showed higher sensitivity and specificity compared to endpoint PCR or serological methods (Gachon et al. 2004; Jeong et al. 2014; Bluth and Bluth 2018). Less starting material was needed, which could then be quantified (Gachon et al. 2004; Jeong et al. 2014; Pallás et al. 2018). Although instrumentation was initially expensive and technical expertise was required, technological advancements and increasing widespread proficiency reduced the impact of these limitations on the use of qPCR (Jeong et al. 2014; Selvarajan and Balasubramanian 2016). The increase in genome data available made the designing of PCR primers easier and more accessible (Schaad and Frederick 2002; Gachon et al. 2004; Jeong et al. 2014).

Within the modern qPCR approach, differing modes of generating the measured fluorescent signal have been developed. These can be grouped into probe-based methods, for example TaqMan[TM] and molecular beacons (Figure 1.20A), or non-probe methods using intercalating fluorescent dyes such as EtBr and SYBR Green (Figure 1.20B; Schaad and Frederick 2002; Gachon et al. 2004; Boonham et al. 2014; Bluth and Bluth 2018).

In general, probe-based methods rely on the hybridisation of oligonucleotide probes in addition to the PCR primers. These probes contain a fluorophore and a fluorophore quenching

molecule (Figure 1.20A). After the probe has bound to the denatured DNA, the fluorophore and quencher molecule become separated, or the fluorophore is released from the probe by the action of the *taq* DNA polymerase, causing the fluorophore to fluoresce. The precise mechanisms for each manufacturer of probes or molecular beacons can vary, however, the detection and measurement of fluorescence is consistent (Schaad and Frederick 2002; Gachon et al. 2004; Boonham et al. 2014).

In non-probe methods, the fluorescence is generated by fluorescent dyes. These dyes are present in the reaction, but only fluoresce after they have bound to dsDNA (Figure 1.20B; Schaad and Frederick 2002; Gachon et al. 2004; Boonham et al. 2014). Various fluorescent dyes have been developed, seeking to increase fluorescence intensity, resolution and sensitivity, and to decrease any PCR inhibition or preferential binding effects (Monis et al. 2005; Gudnason et al. 2007; Mao et al. 2007; Quanta Bioscience 2010).
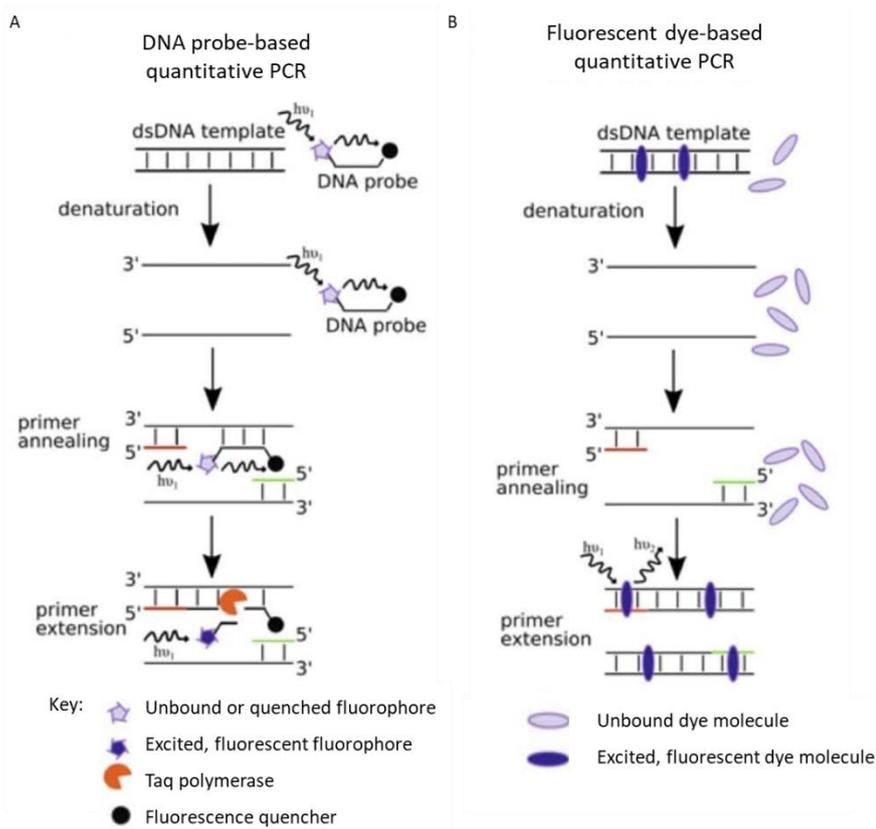
Figure 1.20 Illustration of the two main modes of generating fluorescence for qPCR. A) Illustrates the process for fluorescence generated from probe-based methods. The illustration depicts the TaqMan probe method, where *Taq* polymerase cleaves the fluorophore from the probe and the quencher. B) Illustrates fluorescent dye-based methods, where the dye molecules fluoresce after intercalating with dsDNA (Modified from Nair 2016).

The two qPCR fluorescence modes have different benefits and limitations that impact their use. In general, probe-based methods are more expensive as they require additional time and labour initially to develop target sequence specific primers and probes (Cheng et al. 2013; Dobhal et al. 2016). Whereas fluorescent dye-based methods require only the design of target specific primers, and therefore reduced time and labour costs (Chomič et al. 2011; Cheng et al. 2013; Dobhal et al. 2016). As with endpoint or 'classical' PCR, there have been numerous modifications to the basic qPCR process, including reverse transcriptase qPCR (RT-qPCR) for amplifying RNA targets, and multiplex analysis for amplifying multiple targets in one reaction (Schaad and Frederick 2002; Gachon et al. 2004; Jeong et al. 2014).

The ability of the two fluorescence modes to be used in a multiplex setting has impacted their use. Multiplex detection using probe-based methods is limited due to instrument compatibility and ability to accurately distinguish the fluorescent signals from multiple fluorophores and quenchers (Boonham et al. 2014; Pallás et al. 2018). Using dye-based methods, there is a similar issue if attempting to multiplex with different dyes as their fluorescence spectra overlap, such as the SYBR Green and EvaGreen dyes (Pallás et al. 2018). However, the development of a specific post-qPCR process from fluorescent dye-based methods – melt curve analysis (MCA) – has provided for accurate multiplexing (Varga and James 2006; Cheng et al. 2013; Nolan et al. 2013; Pallás et al. 2018).

MCA is performed by denaturing the PCR product through gradually increasing the temperature. This causes a decrease in fluorescence as the PCR product becomes single-stranded. Monitoring and software analysis of the decreasing fluorescence generates a specific melt curve (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Bluth and Bluth 2018). The melt curve is specific to each target DNA sequence due to the particular melting properties of the specific target sequence (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Bluth and Bluth 2018).

MCA has further progressed into High Resolution Melting (HRM; Section 1.2.4.2.5). Where the use of high-intensity optical detection and fine thermal resolution instruments with dedicated software (Nolan et al. 2013) can construct precise melt curves that can detect genetic variation and mutations, such as single nucleotide polymorphisms (SNPs; Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013).

The core feature of qPCR – the real-time, accurate, quantitative measurement of the target DNA – is ideal for to two main types of applications in plants; gene expression studies, and the detection and quantification of foreign DNA, such as pathogens (Gachon et al. 2004; Fox and

Mumford 2017). Recently, gene expression studies for LNYV and its subgroups was performed using RT-qPCR (Hull 2019). While it was a small scale study, gene expression profiles were shown to be different for the two subgroups across the genes studied (Hull 2019). Although there appear to be no published reports of LNYV detection using an RT-qPCR approach, this method for plant virus detection has remained popular and is well documented (Gachon et al. 2004; Boonham et al. 2014; Jeong et al. 2014; Fox and Mumford 2017). The development of qPCR likely contributed to the technological shift seen in Figure 1.14. Where from 1995 onwards, there was a jump from a 10 % contribution to 30 % contribution of first detections using molecular technologies (Fox and Mumford 2017; Pallás et al. 2018).

1.2.4.2.5 *High Resolution Melting (HRM)*

Following the conclusion of qPCR amplification, MCA or HRM can be performed to distinguish the PCR products. HRM differs from MCA by using brighter fluorescent dyes at high concentrations. It also requires instruments that have been optimised with high-intensity, high-speed optical detection systems for constant capture of the fluorescence along an accurate and finer temperature gradient. This is followed by dedicated software analysis for HRM (Gachon et al. 2004; Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013).

HRM involves slowly heating, and therefore melting, the PCR product. As the PCR product denatures, the intercalating fluorescent dye is slowly released back into the solution and fluorescence decreases. This occurs due to the fluorescent dye only binding and fluorescing with dsDNA (Figure 1.20). As the PCR product is melted into a single-strand state there will be a sharp decrease in fluorescence, which indicates the halfway point between the double-stranded and the single-stranded states of the PCR product. This also indicates the melting temperature (T_m) of the specific PCR product (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Bluth and Bluth 2018).

Using the dedicated software, the T_m point is calculated from a melt curve (Figure 1.21A, and Figure 1.21B) and is then plotted as a melt peak (Figure 1.21C). The melt curve shows the gradual and then sudden decrease in fluorescence as the temperature increases and the PCR product denatures. This data is then normalised to aid in the discrimination of the melt profiles and to eliminate any background fluorescence (Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012).

From the raw data stable pre- and post- melt fluorescence intensity regions are selected (Figure 1.21A). The pre-melt region is used to designate 100 % fluorescence, where every PCR

product is still double-stranded (Applied Biosystems 2009). This point is normalised to a relative fluorescence value, usually 1.0 (Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012). The post-melt region designates where every PCR product is single-stranded and the relative fluorescence value is therefore 0.0 (Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012). The region between the pre- and post- melt regions is referred to as the active melt region (Figure 1.21A) and is displayed in the normalised melt curve plot (Figure 1.21B). Using the normalised melt curve data, the negative derivative of the fluorescence is plotted against the temperature derivative (-dF/dT) and generates a peak shape (Figure 1.21C; Roche Diagnostics 2012). The peak shape aids in the discrimination of any subtle melt profile differences due to sequence variation in the PCR products. This plot is referred to as the normalised melt peak (Figure 1.21C; Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012).
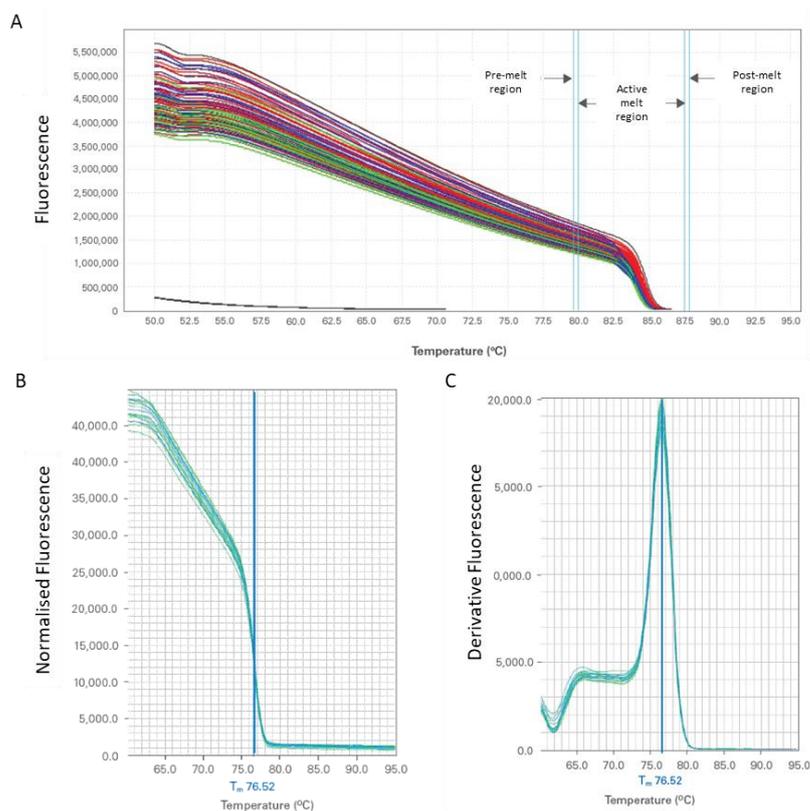
Figure 1.21 High resolution melt data following qPCR. A) A standard HRM melt curve, showing initial high fluorescence as the dye is bound to the double-stranded PCR product. As the temperature increases, the PCR product denatures, releasing the dye and fluorescence decreases. Double blue lines represent the pre- and post-melt regions that are used to normalise the data. Coloured melt curve lines represent different samples. B) Normalised HRM melt curve. The vertical line (bright blue) indicates the halfway point and T_m of the PCR product. Sample melt curves are coloured turquoise. C) A normalised HRM melt peak formed from the derivative fluorescence of (B). The vertical line (bright blue) indicates the T_m of the PCR product and creates the melt peak (Modified from Applied Biosystems 2009).

Differences in the melt peak profile shape or peak temperature across samples allows for discrimination of the PCR products. The melt peak temperature, which is the T_m of the PCR products, is influenced by the size, GC content, and sequence and strand complementarity of the PCR product (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Bluth and Bluth 2018). Differences in any of these properties can alter the melt peak profile shape or peak temperature, leading to sensitive discrimination between multiple PCR products, or detection of sequence variation in the PCR products (Applied Biosystems 2009; Taylor et al. 2011; Bluth and Bluth 2018).

For further discrimination and sample identification, the melt curve data can also be constructed into a difference plot (Figure 1.22). A control or 'wild type' sample is selected as the baseline, and the remaining sample curves are subtracted from the baseline curve (Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012). This allows for greater visualisation of differences in the melt profile shape, and assists in grouping samples with similar or homozygous melt curve profiles (Roche 2008; Applied Biosystems 2009; Taylor et al. 2011; Roche Diagnostics 2012).
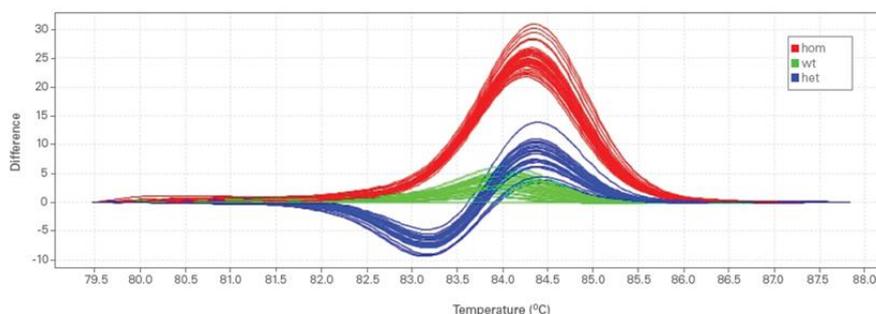
Figure 1.22 Normalised HRM melt curve data displayed as a difference plot to emphasise differences across the samples. Each line represents a separate sample. Sample groupings include homozygous (hom; red), wild type (wt; green) and heterozygous (het; blue). The wild type samples (wt; green) were used as the baseline (Reproduced from Applied Biosystems 2009).

The HRM technique provides a closed tube system that is fast, straightforward, specific, sensitive, and can provide high-throughput analyses (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Bluth and Bluth 2018). It is also non-destructive, allowing further post-PCR analysis of the PCR product, such as gel electrophoresis or sequencing (Taylor et al. 2011; Bluth and Bluth 2018). Key applications of HRM include methylation analysis, genotyping, mutation scanning and SNP detection (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al.

2013; Bluth and Bluth 2018). However, applications of the technology have broadened to include clinical applications (Montgomery et al. 2010), and pathogen detection for human diseases (Diaz and Winchell 2016; Bispo et al. 2018; Shin et al. 2019).

Plant applications using HRM include, genetic fingerprinting, molecular markers, and DNA barcoding (Li et al. 2010; Simko 2016) for plant breeding (Yamagata et al. 2018), authenticity checks of food, wine, seeds (Simko 2016; Pereira et al. 2018) and phytopharmaceuticals (Osathanunkul et al. 2016), and foreign DNA or pathogen detection (Simko 2016). HRM has been used for the detection of a range of plant pathogens, such as fungi infecting pine trees (Luchi et al. 2011) and bananas (Wong et al. 2013), bacteria infecting olive, oleander, and ash trees (Gori et al. 2012), and viruses infecting a range of plants, including vegetables (Cheng et al. 2013; Nie et al. 2016), fruit (Komorowska et al. 2014; Aloisio et al. 2018), and ornamental plants (Dobhal et al. 2016), to list a few here. The use of HRM in plant virus detection is further discussed in Chapter 2.

There has not been a published diagnostic assay for LNYV using RT-qPCR or RT-qPCR followed by HRM. However, an assay using isothermal amplification (Section 1.2.4.2.6) for the detection of LNYV in field samples was developed in 2018 (Fletcher et al.) but required further optimisation. Very recently, Zhang et al. (2020) published the optimised isothermal assay for LNYV detection. The assay used loop-mediated isothermal amplification (LAMP) of nucleic acids to amplify a region of the LNYV N gene.

1.2.4.2.6 *Isothermal Amplification*
A key limitation to PCR technologies for the amplification of nucleic acids was the need for expensive thermocycler instruments and trained technicians with molecular biology skills (Boonham et al. 2014; Selvarajan and Balasubramanian 2016). This influenced the limitation of these technologies to well-funded laboratories and their unsuitability for use in the field directly (Boonham et al. 2014; Fang and Ramasamy 2015; Fox and Mumford 2017). Isothermal amplification was developed as an alternative method to PCR for the amplification of nucleic acids.

Several methods to enable primer binding and extension without the cycling temperature changes associated with denaturation and annealing have been developed. These technologies utilise the thermostable *Bst* DNA polymerase from *Bacillus stearothmophilus* (Aliotta et al. 1996; Notomi et al. 2000; Selvarajan and Balasubramanian 2016), rather than the *taq* polymerase of PCR which operated at high temperatures (Lorenz 2012). However, some of

these showed improved performance after the inclusion of a pre-step requiring a temperature change, or a manual agitation step (Boonham et al. 2014). While other methods required longer reaction times or complicated assay designs (Boonham et al. 2014). The development of LAMP, utilising isothermal auto-cycling strand displacement DNA synthesis (Notomi et al. 2000), has overcome the limitations of other isothermal technologies (Boonham et al. 2014).

Strand displacement amplification (SDA) involves repeated 'bumping' or displacing of the amplified target sequence to allow for a 'new' target strand to be amplified (Walker et al. 1992b; Bluth and Bluth 2018). SDA primers are designed with recognition sites for restriction enzymes, which causes a 'nick' in the amplified strand. Continued sequence extension from the nick caused the initial strand to be displaced. The displaced strand could then be used to amplify the complementary sense of the target sequence, generating exponential amplification (Walker et al. 1992b; Walker et al. 1992a; Bluth and Bluth 2018). The restriction enzyme nick, rather than denaturation and annealing of new primers, meant the reaction could occur at one temperature (Walker et al. 1992b; Walker et al. 1992a; Bluth and Bluth 2018).
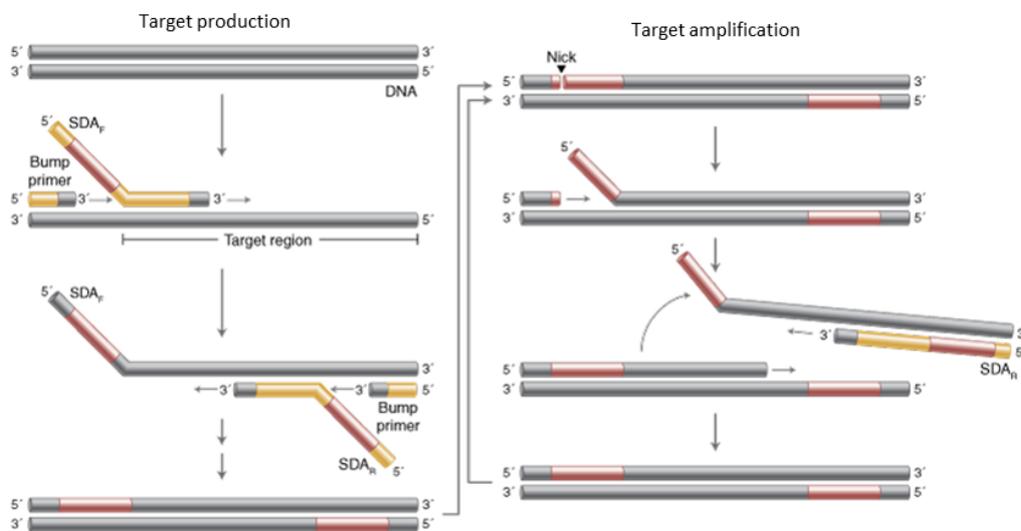
Figure 1.23 Illustration of stand displacement amplification (SDA). Target production shows the binding and extension of the forward SDA primer (SDA$_F$). Binding and extension of the Bump primer upstream from the SDA primer causes the strand to displace. The reverse SDA primer (SDA$_R$) binds to the displaced strand and is extended. A reverse sequence Bump primer also displaces this strand, leading to sequences containing the SDA primers and the target region. Target amplification shows the Nick forming in the SDA primer, allowing a strand to be displaced. A complementary SDA primer binds to the displaced strand and the cycle repeats, amplifying the target region exponentially (Reproduced from New England BioLabs Inc. [NEB] [Date unknown]).

LAMP utilises the concept of SDA, but does not require the use of restriction enzymes to displace the amplified sequences (Notomi et al. 2000). Instead, four to six specially designed primers – internal, external, and more recently, loop primer pairs – are used to generate amplification of both the positive and negative sense of the target sequence (Notomi et al. 2000; Boonham et al. 2014; Selvarajan and Balasubramanian 2016). The use of primers improved the specificity of the assay compared to SDA, as these primers required six to eight specific binding sites on the target DNA (Notomi et al. 2000; Boonham et al. 2014; Selvarajan and Balasubramanian 2016).

The internal primers, forward and reverse (also referred to as forward and backward), each contain two binding sequences; essentially two primers separated by a spacer sequence. These internal primer sequences correspond to the positive and the negative sense sequences of the target DNA (Notomi et al. 2000; Boonham et al. 2014). This allows the 'loop' to form as the internal primer sequences are complementary, whereas the external primers contain only one binding sequence and bind further downstream of the target sequence from the internal primers, essentially acting as the 'bump' primers from the SDA method (Notomi et al. 2000; Boonham et al. 2014).

Complementary amplification of the target DNA is initiated by the appropriate binding sequence of the forward internal primer, and also amplifies the reverse (or backward) primer binding sites (Figure 1.24A, and Figure 1.24B). The binding and subsequent amplification of the external primer, positioned downstream, bumps or displaces the complementary strand (Figure 1.24B). The displaced strand is now single-stranded (Figure 1.24C) and forms a loop at one end due to the 'second' internal primer being complementary, or in the negative sense, of the binding sequence (Figure 1.24D). This stem and loop strand can then be used as the template for the reverse (or backward) internal primers. The reverse primers bind and are extended, amplifying the complementary sequences of the forward internal primer binding sites, before being displaced by the reverse external primer (Figure 1.24D). The displaced reverse strand briefly forms a characteristic dumbbell shape due to loops forming at each end (Figure 1.24E).
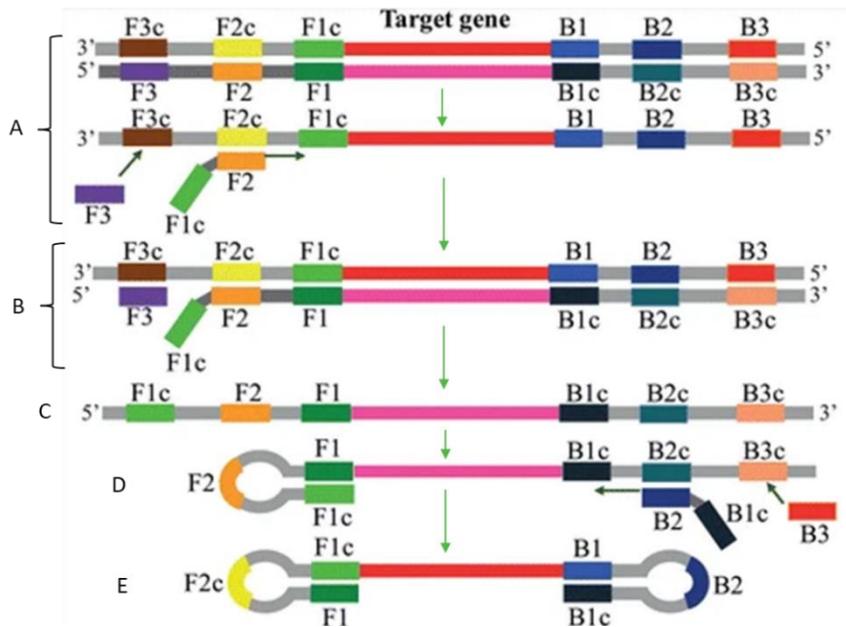
Figure 1.24 Illustration of the initial steps of the isothermal LAMP process generating the starting material. A) Shows the DNA target including the forward (F) and reverse (or backward; B) primer binding sites. The forward internal primer consists of F1c (Forward 1 complement, light green) and F2 (orange), which binds to the F2c (F2 complementary, yellow) sequence and is extended. B) The forward external primer (F3, purple) binds to the complementary sequence downstream F3c (brown) causing strand displacement. C) Shows the displaced strand which also synthesised the reverse (B) primer binding sites. D) The forward loop forms by F1c (light green) binding to F1 (dark green). Figure B and C are repeated with the reverse internal primer, consisting of B1c (black) and B2 (dark blue), and the reverse external primer F3 (red). E) Forward internal primer and the reverse internal primer both form loops from the strand extended by the reverse primers. This forms the characteristic dumbbell structure. LAMP process is continued in Figure 1.25 (Modified from Le and Vu 2017).

The dumbbell structure becomes the starting material for LAMP cycling (Figure 1.25E). From the forward loop self-primed DNA synthesis occurs (Figure 1.25F) while also allowing the binding of the forward internal primer (Figure 1.25G) as this strand was generated by the reverse primers and contains the complementary sequence of the forward primers. After extension at both the new forward internal primer and the self-primed forward primer loop site, the new sequence now contains an additional inverted copy of the target sequence in the stem and a loop at the opposite end via the initial reverse sequence (Figure 1.25H). Self-primed DNA synthesis at the reverse (or backward) loop causes strand displacement (Figure 1.25Ia, and Figure 1.25J), generating one complementary strand of the original stem and loop DNA (Figure 1.25Ia, and Figure 1.25J), and one stem and loop strand with double copies of the target sequence and a loop at the reverse primer sequence site (Figure 1.25Ib). Both of these strand structures (Figure 1.25Ia, and Figure 1.25Ib) serve as template strands for strand displacement by the reverse primers in subsequent cycles (Figure 1.25J, Figure 1.25K, Figure 1.25L, and Figure 1.25M), with each half cycle amplifying the target DNA 3-fold (Notomi et al. 2000).
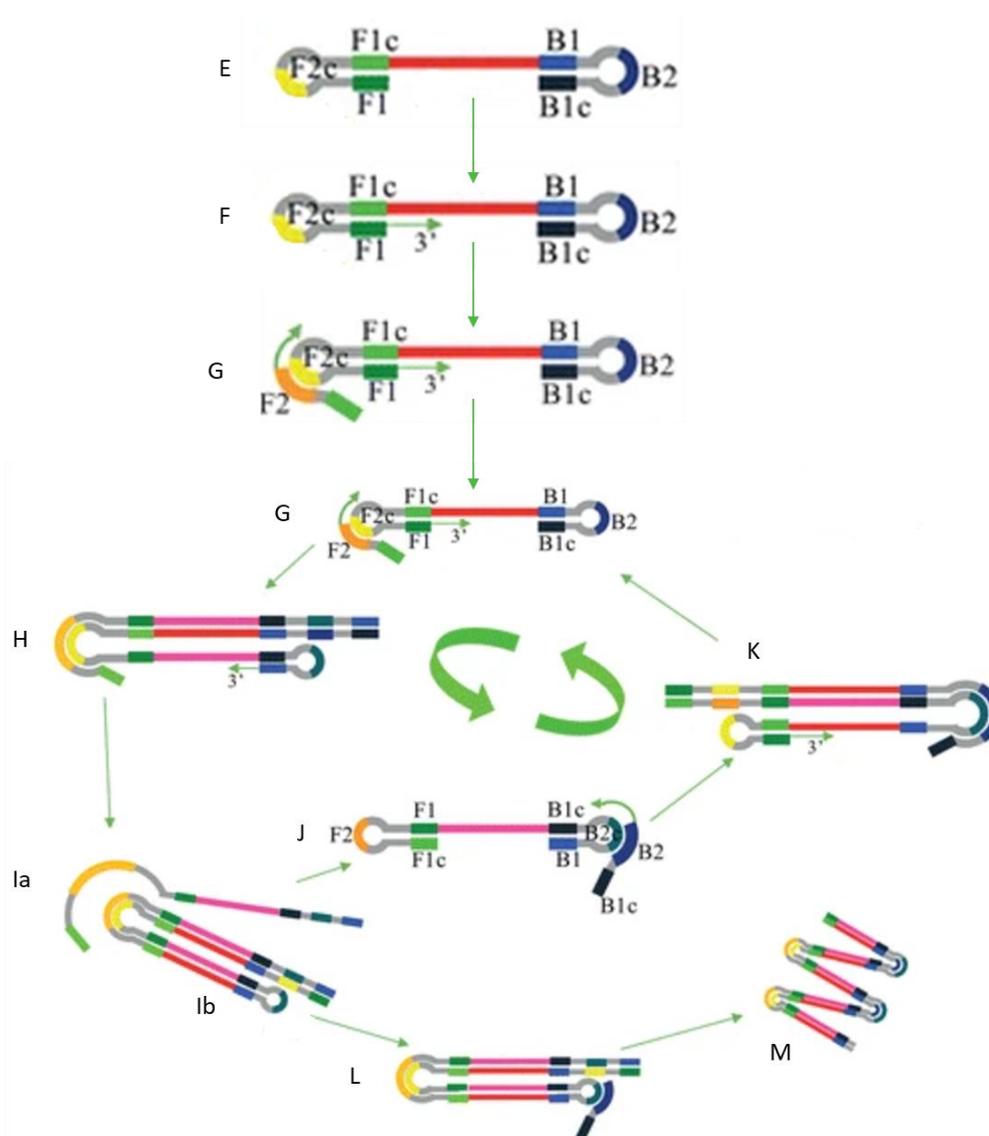
Figure 1.25 Illustration of the cycling process for the isothermal LAMP method. Continued from Figure 1.24. E) Shows characteristic dumbbell structure from Figure 1.24E. F) Shows self-primed DNA synthesis from the internal forward primer (F1, dark green). G) Shows the binding of forward internal primer to the F2c site in the forward loop on the dumbbell structure of Figure E. H) shows the strand now contains an additional inverted copy of the target sequence in the stem and a loop at the opposite end via the initial reverse sequence. Ia) Self primed DNA synthesis at the reverse loop causes strand displacement. J) Displaced strand forms dumbbell structure. K) Shows cycling amplification of the structure from Figure Ia. Ib) Show one stem and loop strand with double copies of the target sequence and a loop at the reverse primer sequence site. L) and M) show amplification of this structure (Modified from Le and Vu 2017).

The resulting amplification produces stem and loop structures of varying length. These are commonly confirmed by size discrimination using gel electrophoresis. This can produce a electrophoresis smear due to the multiple band sizes, or the amplified products can be digested by restriction enzymes prior to electrophoresis (Notomi et al. 2000; Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016). Modern LAMP technologies

use intercalating fluorescent dyes, such as SYBR Green used in qPCR, to detect the amplification by fluorescence or a colour change (Boonham et al. 2014; Selvarajan and Balasubramanian 2016). This modern technique eliminates the time spent running gel electrophoresis, as well as the potential for contamination by opening the sample tubes (Boonham et al. 2014; Selvarajan and Balasubramanian 2016).

The convenience, cost-effective, specific, and rapid nature of isothermal amplification using LAMP makes this technology ideal for on-site, in the field, virus detection (Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016). The isothermal nature of the assay means simpler, cost-effective and more portable equipment can be used. As well as less specialised training is required to perform the assay compared with PCR and qPCR assays (Boonham et al. 2014; Selvarajan and Balasubramanian 2016). LAMP assays proved to be target specific, requiring only a small amount of starting material, while also exhibiting a high tolerance to biological inhibitors or contaminants (Notomi et al. 2000; Boonham et al. 2014; Selvarajan and Balasubramanian 2016).

The LAMP assay can also be modified to detect RNA targets as well by the incorporation of a reverse transcription enzyme to the reaction (RT-LAMP; (Notomi et al. 2000; Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016). LAMP assays, and in particular RT-LAMP assays, are being used widely for the detection of plant disease pathogens (Boonham et al. 2014; Jeong et al. 2014; Selvarajan and Balasubramanian 2016; Le and Vu 2017). A review of publications between 2000 and early 2016 found over 250 peer-reviewed research articles using LAMP assays for the diagnosis of plant diseases (Le and Vu 2017). These publications had documented the use of LAMP technology in the detection of 50 different plant viruses, of which over half utilised RT-LAMP assays (Le and Vu 2017).

The development of an isothermal RT-LAMP assay for LNYV was to enable in the field detection, that was faster and more sensitive compared to the RT-PCR assay commonly used (Fletcher et al. 2018). Zhang et al. (2020) recently published an optimised protocol for the RT-LAMP assay. This included a pair of loop primers enabling detection of a total of eight LNYV specific sequences of the N gene within a region approximately 400 bp in length. This enabled fast, specific and sensitive detection of LNYV. However, ever after optimisation, the RT-LAMP and RT-PCR assays were consistent in detecting LNYV infections in the field samples (Zhang et al. 2020). In addition, although LAMP assays have shown to detect closely related viral species (Le and Vu 2017; Du et al. 2019; Waliullah et al. 2020), there is a lack of evidence around whether this technology is specific enough for detection of viral subgroups or variants within a

species. Often qPCR detection has shown greater sensitivity and specificity over LAMP assays (Boonham et al. 2014; Selvarajan and Balasubramanian 2016; Waliullah et al. 2020).

Recently, Ajithkumar (2018) developed an assay to detect the LNYV subgroups using RT-PCR followed by restriction fragment length polymorphism (RFLP), a common method used for species specific identification (Pallás et al. 2018). Although, Ajithkumar (2018) demonstrated the developed primers were subgroup specific and could be used in a multiplex RT-PCR assay, there were limitations to the assay. The PCR products for each subgroup were of similar size. This meant the PCR products needed to be distinguished by RFLP, or each sample needed to be tested twice – once with the subgroup I primers, and once with the subgroup II primers (Ajithkumar 2018).

Both options for the assay described by Ajithkumar (2018) required additional time, labour and reagent costs (Bluth and Bluth 2018). RFLP involves digesting the PCR products with a restriction enzyme and looking for differences in the fragment lengths to distinguish the two subgroups (Ajithkumar 2018; Bluth and Bluth 2018). While the alternative involves testing each sample twice for greater confidence in a positive LNYV diagnosis and subgroup identification (Ajithkumar 2018). It is likely a subgroup specific RT-LAMP assay would require similar post-assay confirmation or the testing of each sample twice to distinguish the two subgroups.

Ajithkumar (2018) briefly discussed the suitability of the developed LNYV subgroup specific primers, from here on referred to as the Ajithkumar Primers, for use in a multiplex RT-qPCR-HRM assay to diagnose LNYV infection and the LNYV subgroup of samples. An assay of this type would provide a cost- and time- efficient, closed-tube, rapid, specific, sensitive and high throughput diagnostic assay (Applied Biosystems 2009; Taylor et al. 2011; Nolan et al. 2013; Ajithkumar 2018; Pallás et al. 2018).

1.3 Aims and Objectives

LNYV causes lettuce crop losses in NZ and Australia, and the virus can be differentiated into two subgroups (Dietzgen et al. 2007). LNYV subgroup I appears to have died out in Australia with no samples belonging to this subgroup collected since 1993 (Higgins et al. 2016b). It has been suggested that subgroup II has been able to outcompete subgroup I in Australia due to more optimal interactions with insect and/or plant hosts (Higgins et al. 2016b). In contrast, subgroup I was still detected in NZ samples in 2011 and 2017 (Higgins et al. 2016b; Ajithkumar 2018). It has been suggested that recent high crop losses experienced in NZ may indicate a

stronger presence of subgroup II, or perhaps a stronger strain of LNYV has arrived in NZ from Australia (Fletcher et al. 2019). This study attempts to assess these hypotheses and the structure of the NZ LNYV population using the two aims outlined below.

Chapter 2 describes the development of a LNYV subgroup specific assay using RT-qPCR-HRM analysis. The experiments and results described in this Chapter address the following specific objectives:

1) To test the suitability of the Ajithkumar Primers with RT-qPCR-HRM analysis.
2) To test the suitability of various qPCR fluorescent dyes with the primers and LNYV samples.
3) To test the suitability of newly designed primers with qPCR-HRM analysis.
4) To determine the suitability of the developed RT-qPCR-HRM assay in identifying the LNYV subgroups present in a collection of LNYV infected samples.

Chapter 3 describes the phylogenetic analysis of the entire N gene from 43 LNYV isolates from NZ and Australia to further the current understanding of the NZ LNYV population structure. The 43 LNYV isolates includes unpublished NZ LNYV samples, which were subgroup diagnosed as described in Chapter 2. This was achieved by addressing the following specific objectives:

1) Extract and amplify the entire LNYV N gene from unpublished isolates using endpoint RT-PCR.
2) Sequence the unpublished LNYV N gene sequences via Sanger sequencing.
3) Curate the LNYV N gene sequences.
4) Construct ML phylogenetic trees with 1,000 bootstrap replicates using both the nucleotide and translated protein sequences.
5) Analysis of the taxonomic and evolutionary relationships inferred by the phylogenetic trees.

Chapter 2

Development of a Subgroup Specific

Diagnostic Assay for LNYV

Using RT-qPCR-HRM Analysis

2.1 Introduction

There is a need for a quick, sensitive, specific, and cost-effective diagnostic test to detect LNYV and its subgroups. As discussed in Section 1.1, lettuce is a key economic crop in NZ, and recently crop losses have been experienced in the Mid Canterbury and Nelson regions of the South Island, as well as the lower North Island of NZ (Section 1.2.3; Fletcher et al. 2019). These losses were strongly associated with LNYV, a plant virus detected in NZ since 1965 (Fry et al. 1973). It has been suggested that these crop losses could be due to an increased presence of LNYV subgroup II in NZ, or a new strain of the virus from Australia (Fletcher et al. 2019). Virus detection is important for maintaining, and increasing, the actual crop yield (Culliney 2014; Jeong et al. 2014; Savary et al. 2014; Sharma 2014), and would provide a useful tool to monitor the LNYV subgroup population in NZ. To accurately diagnose LNYV infection, and to help determine if recent crop losses in NZ are due to an increased presence of LNYV subgroup II, an assay that could detect between the two subgroups would be invaluable.

2.1.1 LNYV Diagnostic Assays

LNYV detection methods have evolved as new technologies have been developed. As highlighted in Section 1.2.4, initial detection methods involving analysis of symptoms on mechanically inoculated indicator plants and morphology via electron microscopy had limitations. Primarily both methods require extensive specific expertise and experience to accurately detect different viruses (Boonham et al. 2014). The use of *N. glutinosa* as an indicator plant has been used extensively to study LNYV, as mechanical inoculation of lettuce has proved difficult (Stubbs and Grogan 1963; Hull 2019). *N. glutinosa* has proved to be readily susceptible to mechanical sap inoculation (Stubbs and Grogan 1963; Crowley 1967), is a good biological indicator of LNYV infection, and provided a good source of virus particles for purification (Chambers and Francki 1966; Dietzgen et al. 2007). However, symptoms on *N. glutinosa* vary between isolates (Francki et al. 1989), there appears to be no correlation of symptoms and LNYV subgroup or country of origin (NZ or Australian; Higgins et al. 2016) and there is a lack of consistent criteria published on successful conditions for mechanical inoculation (Crowley 1967; Francki et al. 1989; Hull 2019).

The development of target specific technologies such as serology and molecular technologies proved to be faster, permitted the detection of multiple targets or broad spectrum detection, and required less specific expertise and experience to be performed (Section 1.2.4; (Boonham et al. 2014; Jeong et al. 2014; Fox and Mumford 2017). LNYV infection can be detected by

serology and molecular methods such as various ELISA assays, isothermal PCR assay, and end point RT-PCR analysis (Randles and Carver 1971; Francki et al. 1989; Dietzgen et al. 2007; Higgins et al. 2016b; Ajithkumar 2018; Fletcher et al. 2018; Zhang et al. 2020). These technologies were described in detail in Section 1.2.4.

Molecular detection using end point RT-PCR has been used considerably for LNYV detection and have focused on amplifying regions of the LNYV N gene. The first primer pair; LN-1/ LN-2, was designed by Wetzel et al. (1994b), and used by Tomson and Dietzgen (1995) for LNYV detection from crude plant extracts. The primers amplified a 140 bp region in the LNYV N gene (FIG; (Thomson and Dietzgen 1995). Later, additional primer pairs have been developed. The primer pair BCNG1/ BCNG2 was developed to amplify the entire N gene region of 1,500 bp (Callaghan and Dietzgen 2005). While the primer pairs LNYV_440F/ LNYV_1185R, and BCN3/ BCN4, amplify an internal 750 bp, and a 748 bp region, respectively (Figure 2.1; Higgins et al. 2016b).

LNYV subgroup population analysis has involved the use of the BCNG1/ BCNG2 and BCN3/ BCN4 primers for Australian isolates (Callaghan and Dietzgen 2005; Higgins et al. 2016b), and the BCNG1/ BCNG2 and LNYV_440F/ LNYV_1185R primer pairs for NZ isolates (Higgins et al. 2016b; Ajithkumar 2018). Until recently, end point PCR amplification with these primers, followed by sequencing, was the most common method for determining the subgroup of LNYV infected plants (Callaghan and Dietzgen 2005; Higgins et al. 2016b; Ajithkumar 2018).
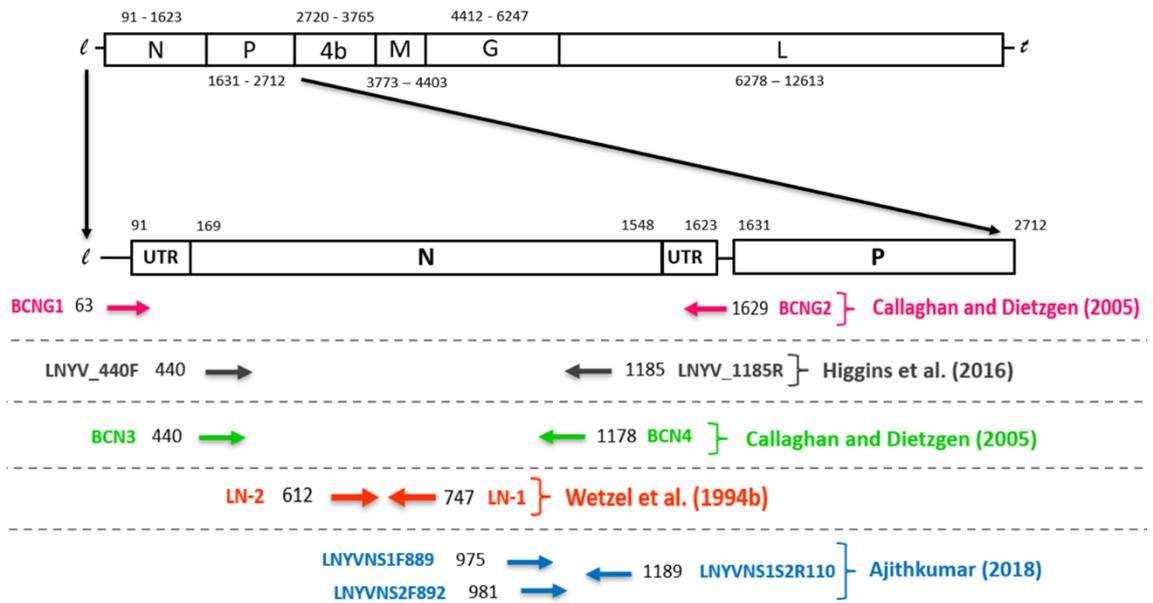
Figure 2.1 LNYV genome organisation illustrating relative primer binding positions. Primer set BCNG1/ BCNG2 (Callaghan and Dietzgen 2005), pink, amplifies the whole 1,500 bp N gene by binding in the intergenic regions on either size of the N gene. Primer set LNYV_440F/ LNYV_1185R (Higgins et al. 2016b), dark grey, amplifies a 750 bp region. Primer set BCN3/ BCN4 (Callaghan and Dietzgen 2005), light green, amplifies a 748 bp region. The LN-1/ LN-2 (Wetzel et al. 1994b), orange, amplifies a 140 bp region. Primer set LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892 (Ajithkumar Primers; Ajithkumar 2018), blue, is subgroup specific and amplifies a 212 bp region of subgroup I and a 209 bp region of subgroup II. Black numbers refer to primer positions on the whole genome. The Ajithkumar Primers were named as per their positions on the N gene sequence alignment (Figure created by author).

2.1.2 Previous Attempts at LNYV Subgroup Detection

An assay that could accurately diagnose LNYV infection and detect between the two subgroups would be invaluable. However, subgroup detection via ELISA assays (Section 1.2.4.1) would require the two subgroups to have detectable distinct epitopes (Jeong et al. 2014), would require a large amount of virus for antiserum production (Fang and Ramasamy 2015), followed by the lengthy virus purification process (Francki et al. 1989). This often makes serological methods not appropriate for closely related viruses (Boonham et al. 2014). In addition, there has been no reported serological difference between different strains of LNYV (Francki et al. 1989).

Viral amplification via molecular PCR methods is highly sensitive and specific, making the detection of closely related, but phenotypically distinct, viruses possible (Section 1.2.4.2; Boonham et al. 2014; Jeong et al. 2014). In 2018, an LNYV subgroup specific diagnostic assay, based on RT-PCR-RFLP was developed by Ajithkumar (2018). Although the Ajithkumar Primers were subgroup specific and showed they could be used in a multiplex assay, they amplified

products of similar size, and had to be distinguished by RFLP using differences in the recognition site for the restriction enzyme *Mae III* (Ajithkumar 2018). This meant each sample needed to be amplified, and the PCR product then digested for subgroup diagnosis. For greater confidence, each sample needed to be tested twice – once with the subgroup I primers, and once with the subgroup II primers – to confirm LNYV diagnosis and the subgroup (Ajithkumar 2018).

The development of an assay based on qPCR would provide a highly-specific, sensitive, flexible and cost-effective diagnostic assay (Section 1.2.4.2.4; Jeong et al. 2014). Although industrial standardisation of reagents has lowered costs, some qPCR detection methods are still more expensive than others (Boonham et al. 2014; Jeong et al. 2014; Aloisio et al. 2018). qPCR analyses based on the use of hybridisation probes requires additional time and labour initially, as probes need to be compatible with the instrument filters, as well as specific to the template. Template specific primers also need to be designed, generating higher costs overall (Cheng et al. 2013; Dobhal et al. 2016). qPCR analyses based on fluorescent binding dyes only require the design of template specific primers, hence reducing experimental design time, labour and costs (Chomič et al. 2011; Cheng et al. 2013; Dobhal et al. 2016). Although these methods may require post-PCR melt analysis, MCA or HRM can be performed by the instrument software. Specific HRM analysis can easily be included at the end of the PCR protocol, adding as little as ten minutes to the programme, and then examined with the instrument software (Chomič et al. 2011; Roche Diagnostics 2012; Cheng et al. 2013; Dobhal et al. 2016).

Analysis using qPCR-HRM has been used for the detection of a range of plant pathogens (Section 1.2.4.2.5). For plant virus detection, both singleplex and multiplex qPCR-HRM assays have been successful and are presented in Table 2.1**.** These articles support this technology as suitable for rapid, sensitive, specific, and cost-effective plant virus detection. These qualities support Ajithkumar's (2018) suggestion that a diagnostic assay using RT-qPCR-HRM could be appropriate for LNYV subgroup detection.

Table 2.1 Recent literature using qPCR-HRM analysis or qPCR-MCA to detect plant viruses or viroids.

Method Used	Plant Virus or Viroid Detected	Reference
Singleplex Analysis		
RT-qPCR-HRM	rose rosette virus (RRV)	(Dobhal et al. 2016)
	grapevine leafroll-associated virus 3 (GLRaV-3) groups: I, II, III, VI	(Bester et al. 2012)
Two-step RT-qPCR-HRM	HSVd variants: HSVd-a, HSVd-b, HSVd-h	(Loconsole et al. 2013)
	potato virus Y (PVY) resistance molecular marker: RY_{sto}	(Nie et al. 2016)
Multiplex Analysis		
RT-qPCR-MCA	plum pox virus (PPV) strains: C, EA, W	(Varga and James 2006)
Two-step RT-qPCR-HRM	cherry necrotic rusty mottle virus (CNRMV) cherry green ring mottle virus (CGRMV)	(Komorowska et al. 2014)
	grapevine fanleaf virus (GFLV) Arabis mosaic virus (ArMV) grapevine leafroll-associated virus 1 (GLRaV-1) GLRaV-3	(Aloisio et al. 2018)
	potato leafroll virus (PLRV) potato virus X (PVX) potato virus A (PVA) potato virus S (PVS) PVY	(Cheng et al. 2013)

As suggested by Ajithkumar (2018), the Ajithkumar Primers may be useful in a RT-qPCR-HRM approach to diagnose the subgroup of samples in one assay. These primers and their amplified products fit the general design criteria for qPCR-HRM analysis (Taylor et al. 2011). The primers anneal at a temperature > 60 °C, which helps ensure specific primer binding and decreases the chance of amplifying non-specific products (Promega Resources [Date unavailable]; Lorenz 2012). The amplified products are < 300 bp in length, and this enhances the likelihood of sensitive detection, and decreases the possibility of multi-phase melting behaviour (Roche 2008; Applied Biosystems 2009; Bustin et al. 2009; Taylor et al. 2011). The Ajithkumar Primers consist of a reverse primer that is non-subgroup specific (LNYVNS1S2R1100) and two subgroup specific forward primers (LNYVNS1F889/ LNYVNS2F892). To provide sufficient specificity of the primers to each subgroup, the forward primers exhibit six nucleotide differences between the subgroups, and each subgroup specific primer pair were vigorously tested in end point RT-PCR reactions to confirm their specificity (Ajithkumar 2018). The subgroup I primer set amplifies a PCR product of 212 bp and the subgroup II primer set amplifies a PCR product of 209 bp (Figure 2.1). These primers were also shown to be suitable for multiplex RT-PCR analysis

(Ajithkumar 2018). It would appear these primers could be suitable for use in multiplex RT-qPCR-HRM analysis for LNYV subgroup diagnosis. If this is not the case, then other primers specific for this type of analysis would need to be designed.

2.1.3 Aims and Objectives

This Chapter describes the development of an assay to enable the diagnosis of LNYV subgroup specific infection of lettuce using RT-qPCR-HRM analysis. The development of this assay will involve the use of samples of known subgroup, after which it will be used to diagnose the subgroup of samples previously untyped but confirmed to be infected with LNYV. The experiments and results described in this Chapter address the following specific objectives:

1) To test the suitability of the Ajithkumar Primers with RT-qPCR-HRM analysis.
2) To test the suitability of various qPCR fluorescent dyes with the primers and samples.
3) To test the suitability of newly designed primers with qPCR-HRM analysis.
4) To determine the suitability of the developed RT-qPCR-HRM assay in identifying the LNYV subgroups present in a collection of LNYV infected samples.

2.2 Materials and Methods

2.2.1 Plant Materials

Leaf material from four plants known to be infected by LNYV was provided by Colleen Higgins (Auckland University of Technology). Lettuce samples collected in 2011 from a farm in Harrisville, Auckland (Higgins et al. 2016b), that had been subgroup typed previously (Ajithkumar 2018) were used as positive controls for experiments. Isolates from Harrisville were designated HV followed by the isolate number. Isolates HV27 (designated NZ2 in Higgins et al. 2016) and HV14 were used as subgroup I specific positive controls, while isolates HV19 (designated NZ1 in Higgins et al. 2016) and HV18 were used as subgroup II specific positive controls. Isolates HV27, HV14, and HV18 were from original lettuce samples. Isolate HV19 was from infected *N. glutinosa* previously created by Ajithkumar (2018), when this isolate has been used from infected *N. glutinosa* it has been specified. Uninfected lettuce and/or *N. glutinosa* were included as negative controls.

Lettuce samples positive for LNYV, but of unknown subgroup, were used for testing the developed RT-qPCR-HRM assay. These were provided by John Fletcher (formerly of The Institute for Plant & Food Research) from various lettuce farms in both the North and South Island of NZ, collected in 2018. The locations and names of the samples used in this analysis are illustrated in Figure 2.2.
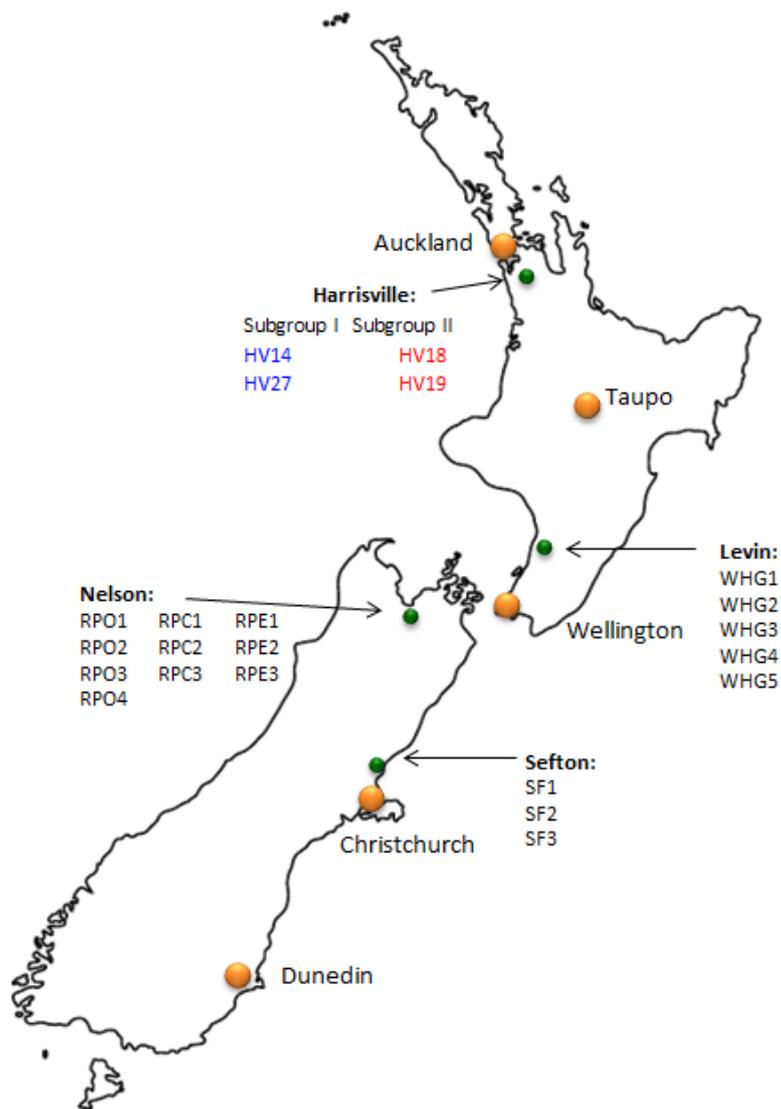
Figure 2.2 Map of New Zealand illustrating the locations of samples used in this Chapter. Orange pins indicate major NZ cities, and green pins indicate sample sites. Blue text represents isolates that have been diagnosed as LNYV subgroup I and red text represents isolates that have been diagnosed as LNYV subgroup II (Higgins et al. 2016b; Ajithkumar 2018). Black text represents isolates that have been diagnosed with LNYV but have not been subgroup typed (Figure created by author).

Isolates from the Nelson region were designated RP, for Richmond Plains, followed by the specific farm identifier (-O, -C, -E) and the isolate number. This naming structure was also used for the isolates from the Levin (WHG) and Sefton (SF) regions.

2.2.2 Experimental Design

Figure 2.3 shows a flow summary of the analysis carried out in this Chapter for the development of a RT-qPCR-HRM assay to diagnose the two LNYV subgroups.
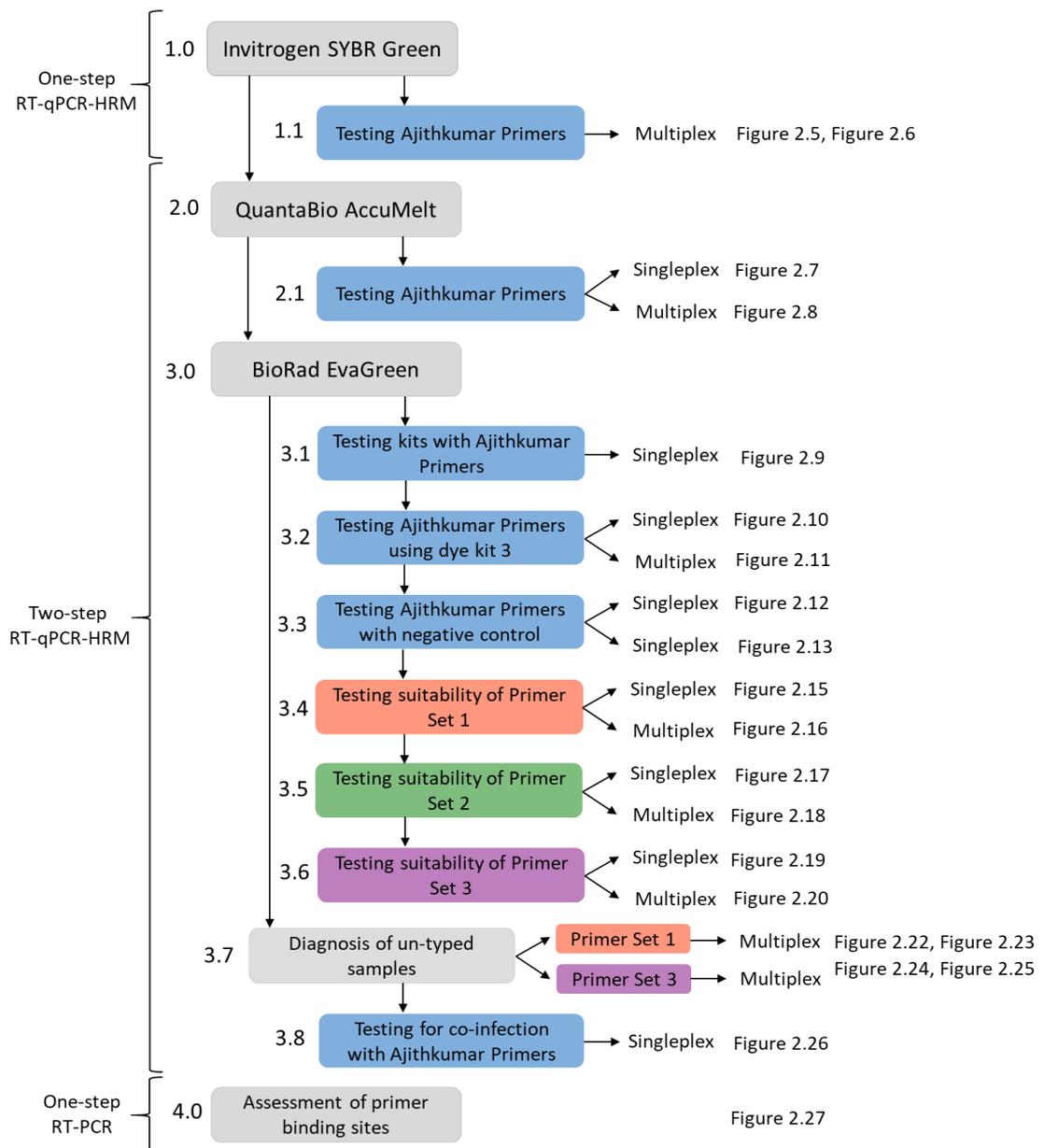
Figure 2.3 Experimental design summary of the analyses carried out in this Chapter to develop a RT-qPCR-HRM assay to diagnose the two LNYV subgroups. Brackets on left indicate assay method. Black arrows indicate sequential flow of experiments. Experiments are numbered for easier reference. Experiment colours relate to primers (where possible) as depicted in Figure 2.1 and Figure 2.14.

2.2.3 RNA Extraction

Total RNA was extracted from frozen plant leaf material using a Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, USA). Using approximately 100 mg of frozen tissue, this was ground via a sterile mortar and pestle with liquid nitrogen until it formed a fine powder. Lysis Solution (500 μL) with pre-added 2-mercaptoethanol (2-ME, 10 μL) was added and the plant material ground further until reaching a liquid state with no obvious pieces of tissue. This liquid was pipetted into a 2 mL collection tube and vortex for 30 seconds, followed by 3 minutes in the

centrifuge (Eppendorf 5430R Centrifuge, New South Wales, Australia) at 15,000 RCF. Lysate was then carefully pipetted into a prepared collection tube with the kit provided blue filtration column. This was then centrifuged for 1 minute at 15,000 RCF and then the blue filtration column was discarded. The Binding Solution (500 µL) was then added to the lysate and mixed via pipette. Up to 700 µL at a time of the combined lysate was added to a prepared collection tube with the kit provided red binding column and centrifuged at 15,000 RCF for 1 minute. The flow through liquid was then decanted. Wash Solution 1 (500 µL) was added to the red binding column, and centrifuged at 15,000 RCF for 1 minute, and the flow through liquid decanted. Then 500 µL of Wash Solution 2 was added and the sample centrifuged at 15,000 for 30 seconds and the flow through liquid decanted. This step was then repeated. After decanting the second flow through liquid of Wash Solution 2, the sample was centrifuged at 15,000 RCF for a further 1 minute. The red binding column was transferred to a new collection tube and 25 µL of Elution Solution was added and incubated at room temperature for 1 minute. The sample was then centrifuged at 15,000 RCF for 1 minute. Elution was repeated with fresh Elution Solution, providing an overall volume of 50 µL of purified total RNA. The total RNA was then stored at - 80 °C until required.

2.2.3.1   Total RNA Quality Testing
The quality, integrity, and concentration of the extracted RNA were assessed via absorbency measurement using a NanoVue spectrometer (GE Healthcare, Life Sciences) and agarose gel electrophoresis.

2.2.3.2   Spectrophotometry
Spectrophotometry was performed using the NanoVue (GE Healthcare, Life Sciences), total RNA protocol. After setting the reference with the RNA extraction kit Elution Solution, 2 µL of total RNA was pipetted onto the reader. The optical density was measured at specific wavelengths 230 nm, 260 nm and 280 nm. These measurements were recorded for each sample.

2.2.3.3   Gel Electrophoresis
The integrity of the extracted RNA was assessed on a 1 % agarose/1 x TBE gel stained with 0.5 µL of 10 µg/mL EtBr. The gel was electrophoresed at 50 Volts for 50 minutes in 1 x TBE buffer

in a Mini-sub® Cell GT Cell Gel tank (Bio-Rad, Auckland, New Zealand). A 100 bp DNA ladder (Solis BioDyne, Estonia) at 2 μL of 0.1 μg/μL was used as a size marker. Each sample well contained 3 μL of purified total RNA with 3 μL of loading dye. The gel was visualised using UV trans-luminescence with an Alpha Imager (Protein Simple, California, USA).

2.2.4 RT-qPCR-HRM Analysis

2.2.4.1   One-Step RT-qPCR-HRM Analysis with Invitrogen SYBR Green Dye

Using a combination of two Invitrogen (now ThermoFisher Scientific) qPCR kits, the LNYV subgroup specific Ajithkumar Primers (LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892), were multiplexed in a one-step RT-qPCR-HRM reaction. Samples were pipetted into a white 96-well plate for analysis in the Roche LightCycler 96 System (Roche Diagnostics, New Zealand). Kit components were thawed and mixed gently via vortex before use. RT-qPCR was prepared for a 12.5 μL reaction with 6.25 μL of SYBR Green UDG mix (from the Invitrogen Platinum SYBR Green qPCR Superscript Mix-UDG kit), 0.5 μL Superscript III (from the Invitrogen Superscript III Platinum One-step qRT-PCR System kit), and 0.25 μL each of the three Ajithkumar Primers (LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892). Each reaction contained 300 ng of total RNA and the remaining reaction was made to a total of 12.5 μL reaction with nuclease free water. Samples consisted of two technical replicates of HV27, a known subgroup I isolate, and HV19, a known subgroup II isolate from infected *N. glutinosa*. All samples were prepared on ice. As the Ajithkumar Primers had already been tested and confirmed not to amplify a product from uninfected lettuce samples (Ajithkumar 2018), only a no template control (NTC) sample was used to detect any possible contamination.

RT-qPCR amplification was carried out using a Roche LightCycler 96 System (Roche Diagnostics, New Zealand), and consisted of the following conditions; cDNA synthesis was carried out by one cycle at 55 °C for 30 minutes, which was followed by amplification with one cycle at 94 °C for 2 minutes, 30 cycles consisting of 15 seconds at 94 °C, 30 seconds at 65 °C and 30 seconds at 68 °C, concluding with a final extension for 5 minutes at 68 °C. HRM analysis was performed as per default settings; one cycle of 60 seconds at 95 °C with ramp rate of 4.4 °C/s, 60 seconds at 40 °C with a ramp rate of 2.2 °C/s, 1 second at 65 °C with a ramp rate of 2.2 °C/s and 1 second at 97 °C with continuous 15 readings/°C.

Analysis was then viewed with the Roche LightCycler 96 Software. The software was used to view the amplification curve, the normalised melt curve, and the normalised HRM melt peak. The default software settings for HRM were used, including the proportional normalisation

algorithm with default sensitivity settings of 50 %, for melt temperature and curve shape discrimination (Roche Diagnostics 2012).

2.2.4.2   Gel Electrophoresis for One-Step RT-qPCR-HRM Analysis with Invitrogen SYBR Green Dye

To confirm the correct amplification from SYBR Green RT-qPCR-HRM, gel electrophoresis was performed. Loading dye (2 μL) was added to the total volume of the RT-qPCR reactions, and electrophoresed at 75 Volts for 50 minutes on a 1.5 % agarose/1 x TBE gel, made as per Section 2.2.3.3.

2.2.5 Two-Step RT-qPCR-HRM Analysis with QuantaBio AccuMelt and BioRad EvaGreen Dyes

QuantaBio AccuMelt, containing SYTO 9 green fluorescent dye, and BioRad EvaGreen qPCR dyes are saturating intercalating fluorescent dyes designed specifically for use in HRM analysis. These dyes were tested for their suitability with the Ajithkumar Primers (LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892). However, both the QuantaBio AccuMelt and BioRad EvaGreen dyes required a separate cDNA synthesis step before qPCR use. cDNA synthesis was carried out using a QuantaBio qScript Flex cDNA Synthesis Kit, or a Solis BioDyne FIREScript cDNA Synthesis Kit, as described below. qPCR-HRM was then carried out using QuantaBio AccuMelt Dye (Section 2.2.5.3) or BioRad Sso Fast EvaGreen Dye (Sections 2.2.5.4, 2.2.5.5, 2.2.6.1 and 2.2.6.2). All qPCR experiments were carried out using white 96-well plates, on a Roche LightCycler 96 System (Roche Diagnostics, New Zealand).

2.2.5.1   cDNA Synthesis with QuantaBio qScript Flex cDNA Synthesis Kit

Using the QuantaBio qScript Flex cDNA Synthesis Kit (QuantaBio, Massachusetts, USA), cDNA for the RNA samples was created as per the protocol for use with the random primers provided with the kit. Briefly, kit components were thawed and mixed gently via vortex before use. Each reaction, on ice, contained 2 μL of kit provided Random Primers, 300 ng of total RNA and nuclease free water to a total volume of 15 μL. Tubes were then briefly mixed by vortexing and centrifuged for 10 seconds. Samples were incubated in a dry heat block at 65 °C for 5 minutes and then snap chilled on ice. qScript Flex Reaction Mix (5 X, 4 μL) and 1 μL of qScript Reverse Transcriptase was added to each sample, then vortexed gently and centrifuged. Samples were then incubated in a Bibby Scientific ™ Techne ™ TC-512 Gradient Thermal Cycler

(Fisher Scientific, England, UK) for 10 minutes at 25 °C, 45 minutes at 42 °C, and 5 minutes at 85 °C before being held at 4 °C. Multiple cDNA synthesis reactions for the same template were then pooled together into one tube, and stored at -20 °C until required.

2.2.5.2   cDNA Synthesis with Solis BioDyne FIREScript cDNA Synthesis Kit

Solis BioDyne FIREScript cDNA Synthesis Kit (Soils BioDyne, Estonia), without primers, was also used to prepare cDNA. Synthesis was performed as per the provided kit protocol. Kit components were thawed and mixed gently via vortexing before use. Briefly, 300 ng of total RNA was combined with 2 µL of the RT Reaction Premix without primers (10 X), 1.5 µL FIREScript Enzyme Mix and 1 µL of Random Primer reagent from the QuantaBio qScript Flex cDNA Synthesis kit (QuantaBio, Massachusetts, USA). The remaining reaction volume of 20 µL was made up of nuclease free water. Samples were gently mixed and lightly centrifuged before incubation in a Bibby Scientific ™ Techne ™ TC-512 Gradient Thermal Cycler (Fisher Scientific, England, UK) for 10 minutes at 25 °C, 30 minutes at 37 °C, and 5 minutes at 85 °C, before being held at 4 °C. Multiple cDNA synthesis reactions for the same template were then pooled together into one tube, and stored at -20 °C until required.

2.2.5.3   qPCR-HRM Analysis with QuantaBio AccuMelt Dye

The Ajithkumar Primers (2018; LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892) were tested for use in qPCR-HRM diagnosis of the LNYV subgroups. The primers were tested first in a singleplex assay for each subgroup, followed by a multiplex assay. The primers LNYVNS1S2R1100/ LNYVNS1F889 were used for testing the known subgroup I sample, HV27, and the primers LNYVNS1S2R1100/ LNYVNS2F892 were used for testing the known subgroup II sample, HV19 from infected *N. glutinosa*. The multiplex assay combined all three primers (LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892) with each of the known LNYV subgroup samples. Reactions were prepared as a half reaction (10 µL) of the QuantaBio AccuMelt Supermix protocol (QuantaBio, Massachusetts, USA). Singleplex reactions consisted of 5 µL AccuMelt Supermix, 0.5 µL each of 10 µM forward (LNYVNS1F889 or LNYVNS2F892) and 10 µM reverse primer (LNYVNS1S2R1100), 2 µL sample cDNA (as prepared in Section 2.2.5.1) and 2 µL nuclease free water. The NTC sample used 4 µL of nuclease free water with no cDNA. Multiplex reactions consisted of 5 µL AccuMelt Supermix, 0.5 µL each of the two 10 µM forward primers and the 10 µM reverse primer, 2 µL sample cDNA (as prepared in Section 2.2.5.1) and 1.5 µL nuclease free water. The NTC sample consisted of 3.5 µL nuclease free water with no cDNA. As

the LNYV specific primers had already been tested and confirmed not to amplify a product from uninfected lettuce samples (Ajithkumar 2018), only an NTC sample was used for initial experiments. All assays were conducted with five technical replicates. PCR conditions used were as above (Section 2.2.4.1). LightCycler 96 software was used for analysis of the amplification curves and normalised HRM melt peaks. For multiplex assays the difference plot was also examined. The difference plot is calculated by subtracting the normalised curve of the baseline sample from the normalised curves of the other samples (Roche Diagnostics 2012).

2.2.5.4   qPCR-HRM Analysis with BioRad EvaGreen Dye to Test Kit Reagents

Five BioRad SsoFast EvaGreen Supermix kits (Bio-Rad Laboratoies, Inc. New Zealand), each with five tubes of EvaGreen Supermix dye, were available and each were tested. Testing was performed to determine if the kits were still suitable for use, and secondly, to determine if there was any variation between the kits and the tubes within each kit. Each tube of EvaGreen supermix was tested in a singleplex assay with a known LNYV subgroup I sample, HV27, and a known LNYV subgroup II sample, HV19 from infected *N. glutinosa*. The Ajithkumar Primers were used, and the subgroup I primers (LNYVNS1S2R1100/ LNYVNS1F889) were used for testing HV27, and the subgroup II primers (LNYVNS1S2R1100/ LNYVNS2F892) were used for testing the HV19 isolate from infected *N. glutinosa*. A half reaction (10 μL) of the Bio-Rad SsoFast EvaGreen Supermix was set up, consisting of 5 μL EvaGreen Supermix, 0.5 μL each of 10 μM forward and 10 μM reverse primer, 2 μL sample cDNA and 2 μL nuclease free water. The NTC sample consisted of 4 μL nuclease free water with no cDNA. PCR conditions were as described in Section 2.2.4.1. As the LNYV specific primers had already been tested and confirmed not to amplify with uninfected lettuce samples (Ajithkumar 2018), only an NTC sample was included for initial experiments. The qPCR-HRM analysis was then examined with the Roche LightCycler 96 Software, using the amplification curve and normalised HRM melt peak analysis.

2.2.5.5   Testing Ajithkumar Primers with BioRad EvaGreen Dye in Singleplex and Multiplex RT-qPCR-HRM Analysis

After confirming the usefulness of the EvaGreen reagents and dyes in each of the five kits tested, the kit with the best performance was selected for testing of the Ajithkumar Primers for use in RT-qPCR-HRM analysis. The subgroup I specific primers (LNYVNS1S2R1100/ LNYVNS1F889), and the subgroup II specific primers (LNYVNS1S2R1100/ LNYVNS2F892) were used for testing the kit with known subgroup I and subgroup II samples, HV27 and HV19,

respectively. Reactions were made up as described in Section 2.2.5.4 with five technical replicates, and PCR conditions were as described in Section 2.2.4.1. Singleplex reactions contained the subgroup specific primer pairs. While multiplex reactions contained all three primers, with the appropriate reduction in the volume of nuclease free water. Experiments included an NTC, and uninfected lettuce and *N. glutinosa* as negative control samples, or just an NTC. The qPCR analysis was then viewed with the Roche LightCycler 96 software. The software was used to view the amplification curve and normalised HRM melt peaks for both the singleplex and multiplex analysis, and the difference plot was also examined for the multiplex analysis.

2.2.6 Primer Design for Use in Two-Step RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

The Ajithkumar Primers were found to be not appropriate for RT-qPCR-HRM analysis. Therefore, additional primers designed specifically for two-step RT-qPCR-HRM analysis were developed to optimise LNYV subgroup diagnosis. Careful and specific primer design is essential for accurate and sensitive amplification of the desired target region by qPCR (Taylor et al. 2011; Arif and Ochoa-Corona 2013; Jeong et al. 2014). The LNYV N gene sequences were examined, and published general primer design guidelines were followed for primer design (Applied Biosystems 2009; Bustin et al. 2009; Taylor et al. 2011).

The available published LNYV N gene sequences were collected from the National Centre for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) database. NZ and Australian LNYV isolates, where the N gene sequences had been isolated using the BCNG1/ BCNG2 primer pair (Callaghan and Dietzgen 2005), and used in previous studies (Higgins et al. 2016b; Ajithkumar 2018), were analysed in Geneious v6.0.6 (www.geneious.com, Biomatters, Auckland, New Zealand). A total of 27 LNYV N gene sequences were aligned via the MUSCLE (Edgar 2004) algorithm. The aligned sequences were then manually examined for possible primer positions. The Ajithkumar Primers were used as a point of reference. Characteristics of potential primers were analysed via the Integrated DNA Technologies (IDT) Oligo-Analyser software (https://sg.idtdna.com). Parameters were set at the default for qPCR, with DNA as the target type for use with cDNA samples. The primer sequence length, GC content, predicted melt or annealing temperature, and secondary structures (hairpin and self-dimer) were all assessed. Primers were designed to have a length between 15 – 30 nt, a GC content between 40 – 60 % and a predicted melt temperature about 60 °C, with < 5 °C difference in melt temperature between primers (Applied Biosystems 2009; Taylor et al. 2011; Lorenz 2012). Sequences with

potential for secondary structures were avoided (Applied Biosystems 2009; Lorenz 2012). Primer sequences were also analysed using the NCBI Basic Local Alignment Search Tool (BLAST; https://blast.ncbi.nlm.nih.gov/Blast.cgi) as an initial confirmation of LNYV specificity. In addition, the PCR products for each subgroup from each primer pair were also assessed for PCR product length, GC content and predicted melt temperature via the IDT Oligo-Analyser software. Three final primer sets were synthesised by IDT (Integrated DNA Technologies Pte. Ltd. Singapore: https://sg.idtdna.com). Analysing three primer sets provides robust analysis to determine the most optimal primer combination (Bustin et al. 2009; Lorenz 2012). Features of the three final primer sets are shown in Table 2.2**.**

2.2.6.1 Testing the Specificity of the LNYV Primers Developed for RT-qPCR-HRM Analysis

To confirm LNYV virus and LNYV subgroup specificity of the newly developed primer sets, these were tested in singleplex assays of each subgroup primer pair with known LNYV subgroup I sample, HV14, known subgroup II sample, HV18, uninfected lettuce and an NTC, all with five technical replicates. This was followed by multiplex assays for each primer set with HV14, HV18, uninfected lettuce and an NTC with five technical replicates.

Sample cDNA was prepared as per Section 2.2.5.2, and reactions were prepared as per Section 2.2.5.4 for singleplex assays. Multiplex assays were prepared as per Section 2.2.5.5 except for Primer Set 3, which consisted of four primers (two primer pairs for each subgroup), and required reactions to be prepared with 0.5 µL of the 10 µM fourth primer and only 1 µL of nuclease-free water. PCR conditions were as described in Section 2.2.4.1, except for the annealing temperature which was calculated as the average of the predicted melting temperatures for each primer set (Table 2.2). Primer Set 1 annealing temperature was 60 °C, Primer Set 2 annealing temperature was 63 °C and Primer Set 3 annealing temperature was 65 °C.

Table 2.2 Design features of primers developed specifically for qPCR-HRM analysis. Primers were named as per their location in the LNYV N gene sequence alignment used for primer design. Primer sequences consisted of IUPAC codes where two nucleotides were common among sequences in a particular position. The IUPAC codes used were; M representing A or C, R representing G or A, K representing G or T, and Y representing T or C. The annealing temperature was calculated as the average of the predicted T_m's for each primer set.

	Primer name	Primer sequence	Length of primer	GC %	Predicted T_m (°C)	Annealing temperature (°C)
Primer Set 1	LNYVNS1S2F_351	TGA CAC AGA TTC AGA ACA ACT C	22	40.9	60.5	60
	LNYVNS1R_423	GGG ATT GAM ACA GCA AGG	18	55.6	60.2	
	LNYVNS2R_477	TGA TCT GAA GAA GGA GGT GTT A	22	40.9	60.8	
Primer Set 2	LNYVNS1S2F_1082	GAG ATG GAT TGT CCG GCC A	19	57.9	63.8	63
	LNYVNS1R_1339	RTC KAG ATT TTT GAT GGC AAA TAT T	22	50	64.7	
	LNYVNS2R_1224	AGT GCC CCG AAG TAC TTA GGA T	25	32	61.2	
Primer Set 3	LNYVNS1F_844	YAC KGA AGC AGC CTT AGT RTC AT	23	43.5	63.2	65
	LNYVNS1R_916	TAA AGT CCA TGT GGA TAA GAC TTC AGG T	28	39.3	65.6	
	LNYVNS2F_844	GAC AGA GGC AGC TTT GGT GTC AC	23	56.5	67.0	
	LNYVNS2R_916	CAG TGT CCA TGT CGA TAA CAC TTC G	25	48	64.9	

2.2.6.2 LNYV Subgroup Diagnosis with Primer Sets 1 and 3 with BioRad EvaGreen Dye in Two-Step RT-qPCR-HRM Analysis

Primer Sets 1 and 3 had been identified as good candidates for diagnosis of LNYV subgroups using two-step RT-qPCR-HRM analysis. Their ability to distinguish previously untyped LNYV infected samples was tested in respective multiplex qPCR-HRM assays with BioRad EvaGreen dye. The assays consisted of three technical replicates of positive controls, known subgroup I sample, HV14, and known subgroup II sample, HV18, as well as NTC, and 18 lettuce samples confirmed to be infected with LNYV (2018, personal communication from J Fletcher to C.M Higgins), but had not been subgroup typed. The sampling locations for these lettuce samples are illustrated in Figure 2.2. Multiplex reactions were set up and qPCR conditions were as per Section 2.2.6.1 for each respective primer set.

Samples that gave irregular results (RPC1, RPC2, and SF1) were then tested in singleplex reactions with the Ajithkumar Primers as described in Section 2.2.5.4, with three technical replicates. The amplification curves and normalised HRM melt peaks were examined, along with gel electrophoresis analysis (Section 2.2.4.2). Using data from one-step RT-PCR analysis followed by Sanger sequencing, carried out in Chapter 3, the primer binding sites were visually assessed alongside the N gene sequences for these samples and the subgroup positive control samples.

2.3 Results

2.3.1 Testing RNA Quality

Prior to the development of a diagnostic assay the quality and concentration of the extracted RNA must be suitable for downstream analysis. This was determined via spectrophotometry using the NanoVue spectrometer (GE Healthcare, Life Sciences). The A_{260} / A_{280} values for each sample were around the desirable value of 2.0, and within the acceptable range of 1.8 – 2.2. The A_{260} / A_{230} values were > 1.7, which is the acceptable cut off value for RNA (Sigma Aldrich [Date unknown]; Triabassi [Date unknown]; Wieczorek et al. 2012). This indicated that the purified total RNA of the samples was suitable for downstream analysis. The concentration readings from the spectrophotometry were used for the standardisation of samples in downstream analyses (Roche 2008; Applied Biosystems 2009).

The integrity of the RNA was also checked via agarose gel electrophoresis. This was used to determine if the RNA sample had degraded during the extraction process. An RNA sample with good integrity is expected to show two bands; one brighter band at ~ 1,500 bp, displaying intact 28S ribosomal RNA (rRNA), and another at ~ 700 bp, displaying intact 18S rRNA (Sigma Aldrich [Date unknown]; Wieczorek et al. 2012; Tirabassi 2014).

Figure 2.4 shows an example RNA integrity gel of total RNA extracted from samples used in this study. Intact 28S and 18S ribosomal bands, at the expected sizes of 1,500 bp and 700 bp respectively, were seen. The 28S band was brighter than the 18S band, as was desirable, and no unexpected bands or streaking was produced. This indicated these samples were suitable for downstream analysis.
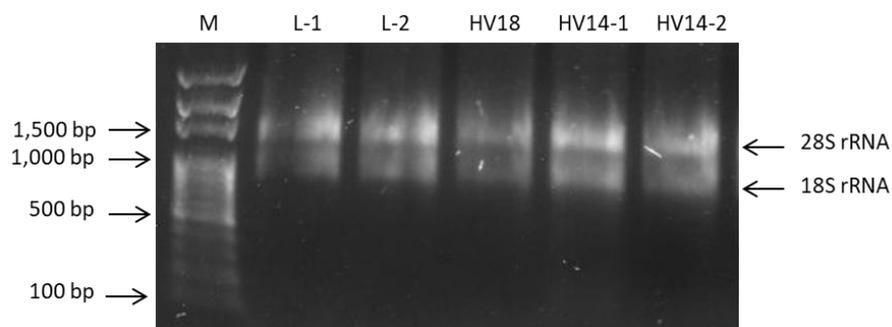
Figure 2.4 Example agarose gel electrophoresis for testing the integrity of total RNA extracted from samples. Lane M = 100 bp ladder. Experimental samples L-1 and L-2 represent uninfected lettuce, while HV18, HV14-1 and HV14-2 represent infected lettuce samples from Harrisville, Auckland.

2.3.2 Testing Ajithkumar Primers in RT-qPCR-HRM Analysis

2.3.2.1   One-Step RT-qPCR-HRM Analysis with Invitrogen SYBR Green Dye

The Ajithkumar Primers were tested for their suitability in a one-step RT-qPCR-HRM assay with Invitrogen SYBR Green dye (Figure 2.5). This experiment represented step 1.1 on Figure 2.3. These primers were able to amplify a PCR product (Figure 2.5A), using RNA from infected lettuce, HV27 (subgroup I, blue), and infected *N. glutinosa,* HV19 (subgroup II, red). The two subgroups could be distinguished, with subgroup II amplification occurring earlier in the qPCR cycles than the subgroup I amplifcation. This was despite equal amounts of total RNA template in each sample reaction. However, the amplification profiles for the technical replicates of each biological sample were inconsistent. No amplification was seen with the NTC replicates.

To assist interpretation and further analysis, the data was normalised using the proportional normalisation algorithm (Section 1.2.4.2.5; Roche Diagnostics 2012). Figure 2.5B shows the melt curves after normalisation of the data. One of the subgroup II technical replicates was normalised to the baseline, while the subgroup I technical replicates produced curve profiles that were again inconsistent.

The HRM melt peaks (Figure 2.5C) were examined to determine the melting temperatures for each PCR product. As this uses the normalised data (Section 1.2.4.2.5) from Figure 2.5B, only three of the four samples produced melt peaks, with one subgroup II technical replicate normalised along the baseline. The single peak for each sample indicated that each reaction amplified a single product. However, broad, discrete melt peaks were observed for each reaction, including the two subgroup I (HV27, blue) technical replicates. The LNYV subgroup II sample (HV19, red) had a melt peak at 81.42 °C, while the subgroup I technical replicates (HV27, blue) produced melt peaks at 79.31 °C and 80.77 °C.

The subgroup I (HV27, blue) technical replicates produced melt peaks with a temperature difference of 1.46 °C between them. As the limit of detection for the LightCycler 96 instrument is 0.2 °C (Roche Diagnostics 2012), a temperature difference greater than this would suggest a true melting difference and not a limitation in the accuracey of temperature dection. This would indicate that the LNYV sequence amplified in each subgroup I technical replicate was different. This was also true for the subgroup II infected samples - if both samples had amplified a PCR product of the same sequence, they would have both produced melt peaks or appeared as flat lines along the baseline for Figure 2.5B and Figure 2.5C. Futhermore, the temperature difference between the two subgroup I replicates was greater than the temperature difference between the closest subgroup I and subgroup II samples. These

observations were unexpected, as each technical replicate came from the same total RNA sample preparation.

To diagnose the LNYV subgroups by HRM, distinct melt peak temperatures and profile shapes is required. The melt peak temperature difference (Figure 2.5C) between the subgroup II (HV19, red) sample and the closest subgroup I (HV27, blue) sample was 0.65 °C. Although this was greater than the instrument limit of detection (0.2 °C; Roche Diagnostics 2012), a temperature difference of at least 1 °C would be needed to be considered acceptable for distinguishing the two subgroups by HRM (Cheng et al. 2013; Osathanunkul et al. 2016). As the HRM profile shapes appeared similar and unexpected results were produced, the samples were not further analysed by difference plot for improved profile shape discrimination.
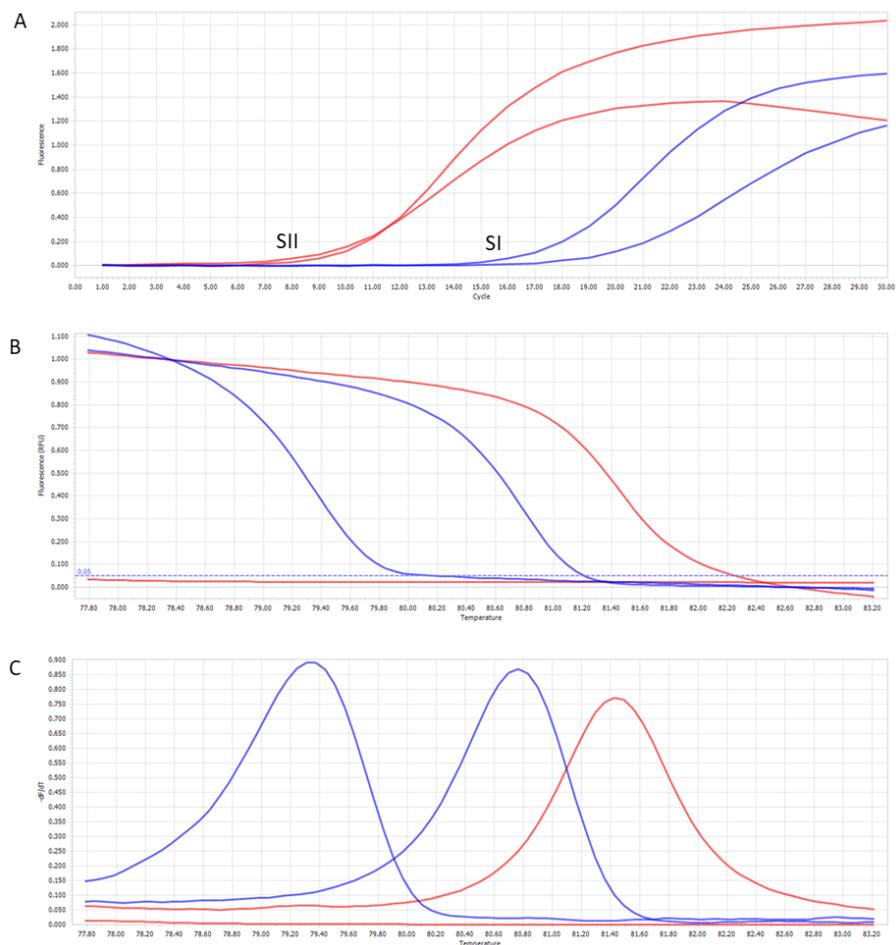
Figure 2.5 Testing the Ajithkumar Primers for LNYV subgroup diagnosis using RT-qPCR-HRM analysis with Invitrogen SYBR Green dye. Subgroup I technical replicates (SI, HV27) are shown in blue, and subgroup II technical replicates (SII, HV19) are shown in red. A) Amplification curve (Cycle number versus fluorescence. B) Normalised melt curve (Temperature versus fluorescence). C) Normalised HRM peaks derived from normalised data shown in B) (Temperature versus -dF/dT).

2.3.2.2 Gel Electrophoresis for One-Step RT-qPCR-HRM Analysis with Invitrogen SYBR Green Dye

To confirm that the PCR products amplified in the multiplex assay described above (Section 2.3.2.1) were of the expected size, size discrimination via gel electrophoresis was carried out (Figure 2.6). Products of the expected ~ 200 bp size (Ajithkumar 2018) for each subgroup and technical replicate were observed. This indicated that the correct PCR amplification occurred in the one-step RT-qPCR-HRM analysis with the Invitrogen SYBR Green dye. Although further confirmation analysis via sequencing could be performed, size discrimination via electrophoresis is a strong indicator of accurate PCR product amplification (Lorenz 2012; Bluth and Bluth 2018). However, while only one band was observed for each sample via gel electrophoresis, the melt curve and HRM melt peak analysis (Figure 2.5B and Figure 2.5C) suggested that the sequences within each PCR product were different.
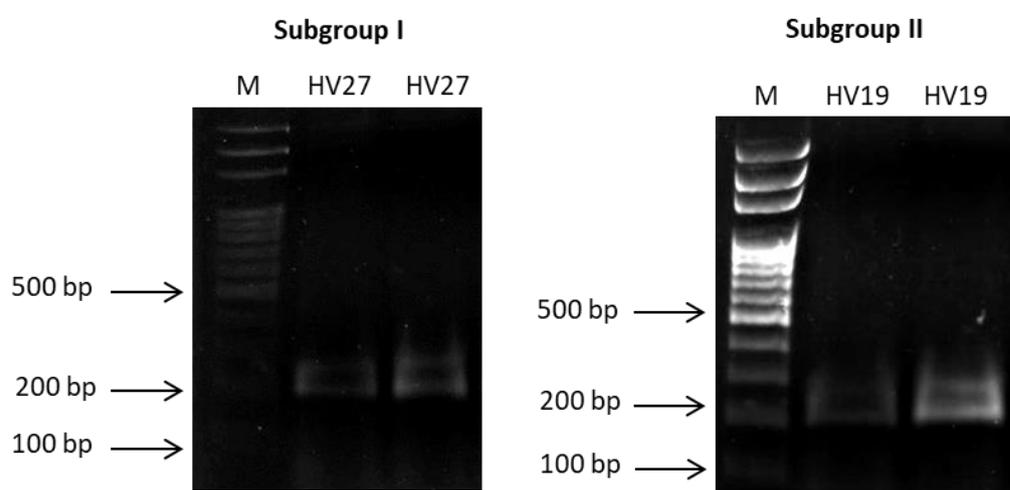
Figure 2.6 Agarose gel electrophoresis for samples analysed in one-step RT-qPCR-HRM assays with SYBR Green dye from Section 2.3.2.1. Expected PCR product at ~ 200 bp can be seen. NTC sample included but showed no amplification of any products (data not shown). Lane M = 100 bp ladder. Subgroup I experimental sample HV27 and technical replicate. Subgroup II sample HV19 from *N. glutinosa* and technical replicate.

At this stage, these results suggested that LNYV subgroup diagnosis by HRM profiles may not be easily distinguishable using the Ajithkumar Primers with SYBR Green dye. SYBR Green dye is a non-saturating intercalating fluorescent dye, and has been reported to be less sensitive in qPCR and melt curve, or HRM, analysis compared with saturating intercalating dyes (Radvanszky et al. 2015; Pereira et al. 2018). To eliminate the dye as the possible source of the less differentiated HRM melt peaks, the Ajithkumar Primers were tested with two different saturating intercalating dyes.

2.3.3 Testing Ajithkumar Primers with QuantaBio AccuMelt and BioRad EvaGreen Dyes

Saturating intercalating fluorescent dyes designed specifically for use in HRM analysis may produce better results than SYBR Green (Sections 2.3.2.1 and 2.3.2.2). QuantaBio's AccuMelt, containing SYTO 9 green fluorescent dye, and BioRad's EvaGreen fluorescent dye, are qPCR dyes designed for use in HRM assays. These dyes were tested by two-step RT-qPCR-HRM analysis for their suitability with the LNYV subgroup specific Ajithkumar Primers (LNYVNS1S2R1100/ LNYVNS1F889/ LNYVNS2F892).

2.3.3.1 Singleplex RT-qPCR-HRM Analysis with QuantaBio AccuMelt Dye

The AccuMelt dye was tested to check its suitability in two-step RT-qPCR-HRM analysis with the Ajithkumar Primers. The primers were tested first in a singleplex assay (Figure 2.7) for each subgroup, followed by a multiplex assay (Figure 2.8). These experiments represented step 2.1 on Figure 2.3.

Singleplex analysis produced amplification of all the technical replicates for subgroup I (HV27, blue; Figure 2.7A), and subgroup II (HV19, red; Figure 2.7B). The amplifcation profiles for both subgroups showed tight, overlapping curves, indicating consistent amplifcation across the replicates for each subgroup. Subgroup II (HV19, red) showed amplification occuring at cycle 14, while amplification of the subgroup I (HV27, blue) replicates occured at cycle 18. Equal amounts of total RNA were included in each reaction, suggesting the earlier amplification of subgroup II (HV19, red) was due to a higher titre of virus within the infected leaves from which the total RNA was extracted. No amplification was seen with the NTC replicates.

The data was normalised and the HRM melt peaks examined (Figure 2.7C and Figure 2.7D). The melt peak profiles indicated amplification of one product in each reaction. For both subgroups, the technical replicates produced some visible variation, creating melt peak ranges, rather than one melt peak temperature. The subgroup I melt peak (HV27, blue; Figure 2.7C) temperatures ranged from 83.24 °C to 83.30 °C, a range which was within the instrument limit of detection (0.2 °C; Roche Diagnostics 2012). The subgroup II (HV19, red; Figure 2.7D) melt peak temperatures ranged from 83.43 °C to 83.57 °C, also within the instrument limit of detection. The temperature difference between the subgroup I temperature range and the subgroup II temperature range was only 0.13 °C, below the minimum acceptable temperature difference of 1 °C (Cheng et al. 2013; Osathanunkul et al. 2016). This suggests the combination of these primers with AccuMelt was inappropriate for subgroup discrimination. However, this

analysis was performed with the primers in singleplex reactions - to confidently analyse the suitability of these primers to diagnose the LNYV subgroups using RT-qPCR-HRM with AccuMelt, multiplex analysis was carried out (Section 2.3.3.2).
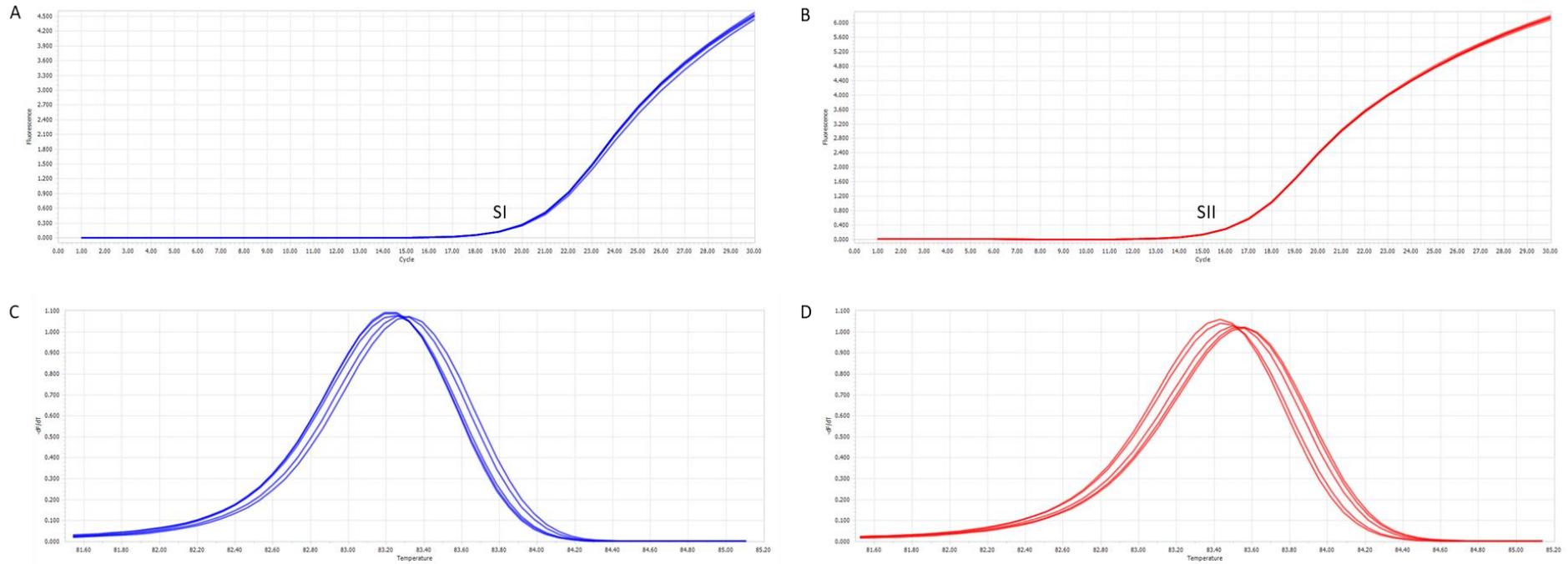
Figure 2.7 Testing the LNYV subgroup specific Ajithkumar Primers in a singleplex assay with QuantaBio's AccuMelt dye. Subgroup I technical replicates (SI, HV27) are shown in blue and subgroup II technical replicates (SII, HV19) are shown in red. A) Subgroup I amplification curve. B) Subgroup II amplification curve. C) Subgroup I normalised HRM melt peaks. D) Subgroup II normalised HRM melt peaks. Showing amplification and HRM melt peaks for all technical replicates of each subgroup analysis. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.3.2   Multiplex RT-qPCR-HRM Analysis with QuantaBio AccuMelt Dye

To confidently assess the suitability of the LNYV subgroup specific Ajithkumar Primers with the AccuMelt dye, the primers were tested in a multiplex assay (Figure 2.8). Amplification was seen for both subgroup samples and all technical replicates (Figure 2.8A). Consistent amplification of the subgroup I and subgroup II samples was indicated by the majority of the amplification curves for each replicate overlapping - except for one subgroup II replicate (HV19, red) which produced a visibly different amplification profile. The multiplex amplification profiles were comparable to the profiles produced when the primers were tested in the singleplex assay (Section 2.3.3.1), again suggesting that LNYV subgroup II (HV19, red) was more abundant in the source tissue than subgroup I (HV27, blue). There was no amplification of the NTC replicates.

The normalised HRM melt peaks for the two subgroups overlapped (Figure 2.8B). This meant that the subgroups could not be discriminated by their melt peak. The melting temperature range for subgroup I (HV27, blue) was 84.06 °C to 84.14 °C while the range for subgroup II (HV19, red) was 84.06 °C to 84.21 °C.

To aid visualisation of any difference between the melt peak profiles, a difference plot was examined (Figure 2.8C). The technical replicates of the subgroup I sample (HV27, blue), were used as the baseline samples, and showed a large degree of variation. Sample replicates did not group near, or on the baseline, as would be expected if the samples had produced identical profiles and melt peaks. The subgroup II (HV19, red) replicates were also spread across the plot, including near the baseline, indicating no substantial difference between the subgroup I and subgroup II samples. These data indicate that the combination of the Ajithkumar Primers with AccuMelt dye was not appropriate for LNYV subgroup diagnosis. The use of a more sensitive, saturating intercalating dye may produce HRM melt peaks with better subgroup discrimination with these primers.
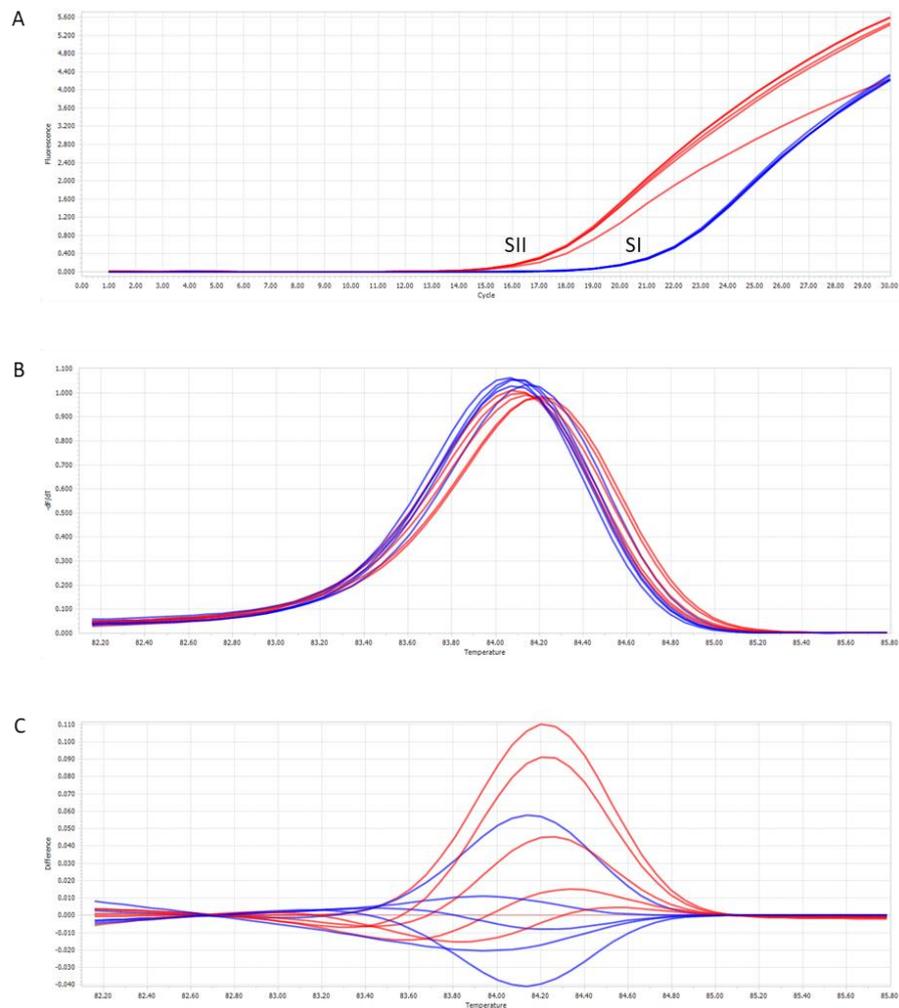
Figure 2.8 Testing the Ajithkumar Primers in a multiplex assay with QuantaBio's AccuMelt dye. Subgroup I technical replicates (SI, HV27) are shown in blue and subgroup II technical replicates (SII, HV19) are shown in red. A) Amplification curve (Cycle number versus fluorescence). B) Normalised HRM melt peaks (Temperature versus -dF/dT). C) Difference plot derived from the normalised HRM met peak data shown in B) with subgroup I technical replicates used as the baseline (Temperature versus difference in -dF/dT). Showing amplification and overlapping HRM profiles for all samples.

2.3.3.3   Testing the Usefulness of BioRad EvaGreen Dye Kits

The BioRad Sso Fast EvaGreen dye was tested as an alternative dye due to its reported higher sensitivity and better performance (Eischeid 2011; Radvanszky et al. 2015). Two-step RT-qPCR-HRM analysis was carried out with the Ajithkumar Primers using the BioRad Sso Fast EvaGreen dye. Five kits of the BioRad Sso Fast EvaGreen dye were available for use that had been in cold storage for some time. To test the usefulness of these kits, each tube from each kit was tested with a known LNYV subgroup I (HV27), and a known subgroup II sample (HV19) in singleplex assays. Testing each tube from each of the five kits acted as technical replicates. In Figure 2.9 the multiple tubes (five) from each kit (numbered 1-5) have been labelled with the same colour for easier comparison. These experiments represented step 3.1 in the experimental design Figure 2.3.

The amplification curves (Figure 2.9A and Figure 2.9B) showed that each tube from each kit amplified a product. This indicates that none of the EvaGreen dye kits had degraded beyond use. For subgroup I (Figure 2.9A) the amplification profiles were consistent across the kits numbered 1, 2, 4 and 5, although some variation across tubes within the kits was visible. Kit 3 (pink) produced a slightly different amplification profile shape, compared to the other kits. Although the amplification profiles for the tubes in this kit were consistent. The amplification curve for subgroup II (Figure 2.9B) showed consistency of the amplification profiles across all the kits, and across the tubes within each kit. There was no amplification of the NTC replicates seen.

The normalised HRM melt peaks (Figure 2.9C and Figure 2.9D) indicated amplification of one product in each reaction. For subgroup I (Figure 2.9C) there was some variation in the melting temperatures of the amplified products, producing a temperature range from 82.53 °C to 82.99 °C. However, the melt peak temperatures and profile shapes appeared consistent for the tubes within each kit. Kits 1 (light green), 2 (orange), 4 (light blue) and 5 (dark blue) produced melt peaks of similar shape and fluorescence level. Kit 3 (pink) produced melting peaks at higher fluorescence, suggestive of more PCR product, with peaks that appeared sharper when compared to the other kits. Kits 1 and 3 produced the lowest within kit melt peak temperature variation of < 0.2 °C.

For subgroup II (Figure 2.9D) variation in the melting temperatures was also seen. This produced a narrow temperature range of 82.92 °C to 83.25 °C for all the tubes in all five kits. Kits 4 (light blue) and 5 (dark blue) amplified products with melting temperature ranges of > 0.2 °C. The melting temperature ranges for kits 1 (light green), 2 (orange), and 3 (pink) were < 0.2 °C. Kit 3 (pink) gave the lowest variation of 0.06 °C.

All five EvaGreen dye kits proved to be usable even after long term storage. Amplification from both LNYV subgroups and their respective primers in singleplex reactions were seen. The kits that produced consistent results with minimal within kit variation was kit 1 (light green) and kit 3 (pink). Kit 3 performed the best overall and was used in subsequent EvaGreen RT-qPCR-HRM assays.
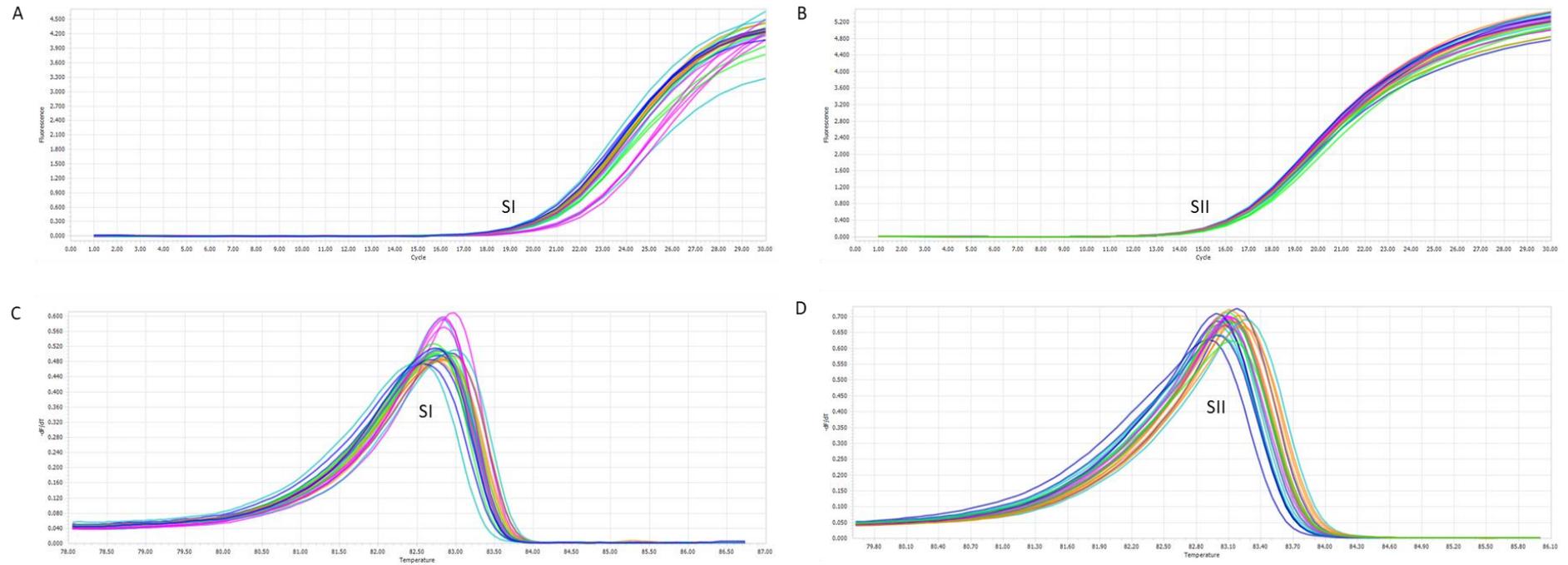
Figure 2.9 Testing the usefulness of BioRad EvaGreen dye kits in singleplex RT-qPCR-HRM analysis with the Ajithkumar Primers. The five tubes from kit 1 are shown in light green, kit 2 are shown in orange, kit 3 are shown in pink, kit 4 are shown in light blue, and kit 5 are shown in dark blue. A) Amplification curve with Ajithkumar subgroup I specific primers and template (SI, HV27). B) Amplification curve with Ajithkumar subgroup II specific primers and template (SII, HV19). C) Normalised HRM melt peak with Ajithkumar subgroup I specific primers and template (SI, HV27). D) Normalised HRM melt peak with Ajithkumar subgroup II specific primers and template (SII, HV19). Shows amplification with all tubes across all five kits, indicating the EvaGreen dye had not degraded. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.3.4 Singleplex RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

In order to better visualise the subgroup HRM profiles further singleplex analysis (Figure 2.10) followed by multiplex analysis (Figure 2.11) with BioRad EvaGreen dye kit 3 was performed. These experiments represented step 3.2 on Figure 2.3.

The singleplex amplification curves (Figure 2.10A and Figure 2.10B) showed amplification across all technical replicates for both subgroup samples. The technical replicates for each LNYV subgroup produced tight, overlapping curves, as was expected. No amplification of the NTC replicates was seen in either singleplex assay.

The normalised HRM melt peaks (Figure 2.10C and Figure 2.10D) indicated amplification of a single product, as expected. The general HRM profile shape for each subgroup was similar; with some visible variation across the technical replicates. For subgroup I (HV27, blue) the melting temperature range was 82.70 °C to 82.83 °C, while subgroup II (HV19, red) it was 83.03 °C to 83.10 °C. The difference between these ranges was 0.2 °C, which is at the limit of detection for the instrument. Although the temperature difference between the two subgroups was larger than that produced in singleplex with the AccuMelt dye (Section 2.3.3.1; Figure 2.7), it was still below the minimum acceptable difference of 1 °C (Cheng et al. 2013; Osathanunkul et al. 2016). As the EvaGreen dye appeared more sensitive than the AccuMelt dye, the primers still showed potential for use in LNYV subgroup diagnostic analysis with RT-qPCR-HRM. Further multiplex analysis was carried out.
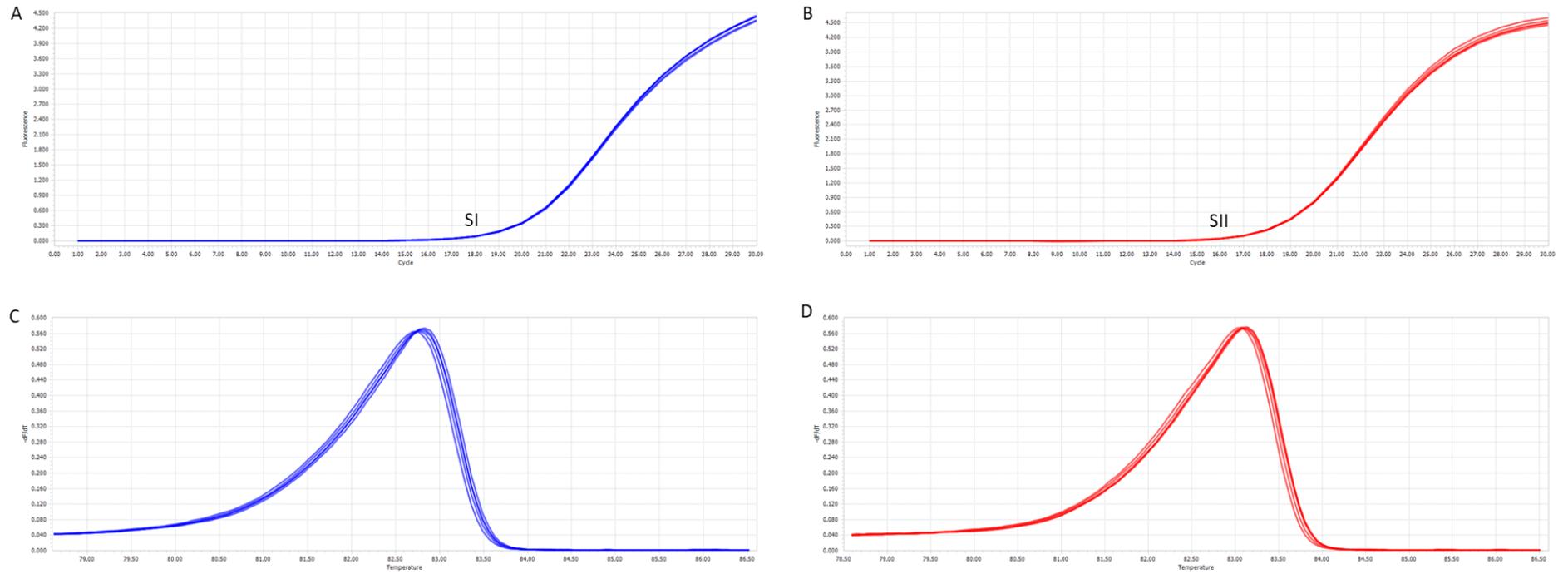
Figure 2.10 Testing the Ajithkumar LNYV subgroup specific primers in a singleplex assay with BioRad EvaGreen dye kit 3. Subgroup I technical replicates (SI, HV27) are shown in blue and subgroup II technical replicates (SII, HV19) are shown in red. A) Subgroup I amplification curve. B) Subgroup II amplification curve. C) Subgroup I normalised HRM melt peaks. D) Subgroup II normalised HRM melt peaks. Showing tight amplification and single melt peaks with both LNYV subgroups. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.3.5   Multiplex RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

Multiplex reactions with the Ajithkumar Primers using the BioRad EvaGreen dye kit 3 was performed to confirm HRM peak profiles of each subgroup in multiplex reactions. A multiplex RT-qPCR-HRM assay for LNYV subgroup diagnosis would reduce time and costs compared to singleplex analysis.

The amplification curve (Figure 2.11A) showed amplification of all five technical replicates for both subgroup I (SI, blue), and subgroup II (SII, red). Two distinct amplification profiles were produced. The amplification curves for the technical replicates of each subgroup overlapped tightly within the qPCR exponential phase. Subgroup II amplified at an earlier cycle, suggesting a higher viral titre compared to subgroup I. No amplification of the NTC replicates was seen.

The normalised HRM melt peak (Figure 2.11B) showed two melt peaks. One melt peak for each LNYV subgroup was seen, indicating amplification of one product in every reaction. Overlap of the two melt peaks was apparent, as was variation among the replicate templates for each subgroup. The subgroup I (HV27, blue) melt peak temperatures ranged from 83.16 °C to 83.23 °C, while the subgroup II (HV19, red) melt peak temperatures ranged from 83.36 °C to 83.43 °C. Both subgroup temperature ranges generated variation of 0.07 °C. However, the temperature difference between the two subgroups, 0.13°C, was below the instrument limit of detection and smaller than that observed in the singleplex assay (Section 2.3.3.4). This would indicate the two subgroups could not be discriminated by their HRM melt peaks.

To allow better examination of the HRM profile shapes and the variation across the technical replicates, the difference plot was assessed (Figure 2.11C). The subgroup I (HV27, blue) technical replicates were used as the baseline samples. The variation across the subgroup I technical replicates was easily seen, and similar variation was also seen across the subgroup II (HV19, red) technical replicates. Although there was clear differences in the profile shapes between the two subgroups, there appeared to be a large amount of variation in the technical replicates for each subgroup, which was undesirable.

Figure 2.11 Testing the Ajithkumar Primers in a multiplex assay with BioRad EvaGreen dye kit 3. Subgroup I technical replicates (SI, HV27) are shown in blue and subgroup II technical replicates (SII, HV19) are shown in red. A) Amplification curve (Cycle number versus fluorescence). B) Normalised HRM melt peaks (Temperature versus -dF/dT). C) Difference plot derived from the normalised HRM met peak data shown in B) with subgroup I technical replicates used as the baseline (Temperature versus difference in -dF/dT). Shows amplification of both subgroups with overlapping HRM profiles and melting curve variation across the technical replicates for each subgroup.

2.3.3.6 Testing the LNYV Specificity of Ajithkumar Primers Using RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

Multiplex analysis of the Ajithkumar Primers with the BioRad EvaGreen dye (Figure 2.11) showed two distinct melt peaks for each subgroup. While, there was variation across the technical replicates for each subgroup and the temperature difference between the two subgroups was below the instruments limit of detection, the results support the possibility of diagnosing the LNYV subgroups using a RT-qPCR-HRM approach. The BioRad Sso Fast EvaGreen dye proved more sensitive than the QuantaBio's AccuMelt dye (Section 2.3.3.2) as evidenced by greater distinction between the two LNYV subgroups when analysed in singleplex and multiplex two-step RT-qPCR-HRM assays. However, the LNYV subgroup specific Ajithkumar

Primers, designed for use in RT-PCR, appeared unable to differentiate the two subgroups in RT-qPCR-HRM analysis.

Using traditional end point RT-PCR analysis, Ajithkumar (2018) had shown that these primers were LNYV subgroup specific. They amplified LNYV from infected leaf tissue, with no non-specific amplification. It had been assumed here that this would also be true in a RT-qPCR-HRM assay, and therefore uninfected samples were not included in the above analyses. As qPCR is more sensitive, and specific, compared to traditional PCR methods (Boonham et al. 2014; Jeong et al. 2014), it was decided to use the BioRad Sso Fast EvaGreen dye kit 3 to confirm that non-specific amplification wasn't influencing the HRM melt peak temperatures. The primers were tested in singleplex two-step RT-qPCR-HRM assays with uninfected lettuce and uninfected *N. glutinosa*, along with the known subgroup I, HV27 from infected lettuce, and known subgroup II, HV19 from infected *N. glutinosa*, representing step 3.3 in the experimental design summary (Figure 2.3).

2.3.3.6.1 *Testing the Ajithkumar Subgroup I Specific Primers with BioRad EvaGreen Dye*

The amplification curve (Figure 2.12) showed amplification with the subgroup I (HV27, blue) technical replicates only. No amplification was seen with the technical replicates for subgroup II (HV19), uninfected lettuce, uninfected *N. glutinosa*, or the NTC. This is consistent with the findings of Ajithkumar (2018) and showed these primers are specific to LNYV subgroup I in a RT-qPCR-HRM assay. As this analysis was implemented to determine if non-specific amplification was occurring from plant RNA, these data were not further analysed by HRM melt peak analysis. These data confirm that the HRM analyses from Sections 2.3.3.4 and 2.3.3.5 were not influenced by non-specific amplification products.
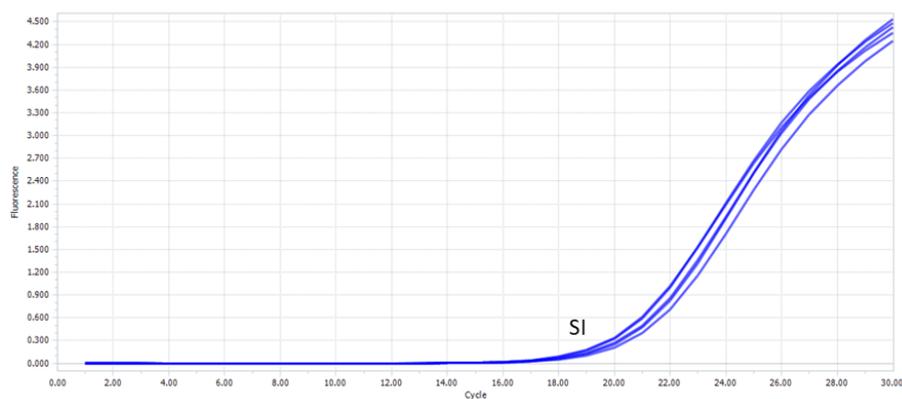
Figure 2.12 Amplification curve for testing the LNYV specificity of Ajithkumar subgroup I primers in a singleplex two-step RT-qPCR-HRM analysis with BioRad EvaGreen dye (Cycle versus fluorescence). Subgroup I (HV27) technical replicates are shown in blue. No amplification of the other samples was seen.

2.3.3.6.2 *Testing the Ajithkumar Subgroup II Specific Primers with BioRad EvaGreen Dye*

It was discovered that the subgroup II (HV19, red) technical replicates and 2 of the 5 uninfected lettuce technical replicates amplified with the Ajithkumar subgroup II specific primers. The amplification curve (Figure 2.13A) showed consistent amplification of the known subgroup II (HV19, red) technical replicates, and there was no amplification of the known subgroup I sample (HV27) or the uninfected *N. glutinosa* technical replicate samples. However, very late amplification from uninfected lettuce (light green) was seen from 2 of the 5 technical replicates, with an increase in fluorescent signal from cycle 28. This amplification was unexpected; however, to rule this out as specific amplification, these data were normalised (Figure 2.13B) and the HRM melt peaks (Figure 2.13C) examined. No amplification of the NTC technical replicates was seen, indicating there was no contamination in the reactions.

Normalisation of the data showed the uninfected lettuce technical replicates normalised to below the baseline. The normalised melt curve (Figure 2.13B) and the HRM melt peaks (Figure 2.13C) revealed the uninfected lettuce technical replicates as a flat light green line. This confirms very low fluorescence was detected for these samples. The subgroup II technical replicates produced an HRM melt peak with a small amount of variation, as was expected. These data suggest the late amplification and fluorescence of the uninfected lettuce was not the result of specific amplification, and was not sufficient to influence the HRM melt peak temperatures (Sections 2.3.3.4 and 2.3.3.5).
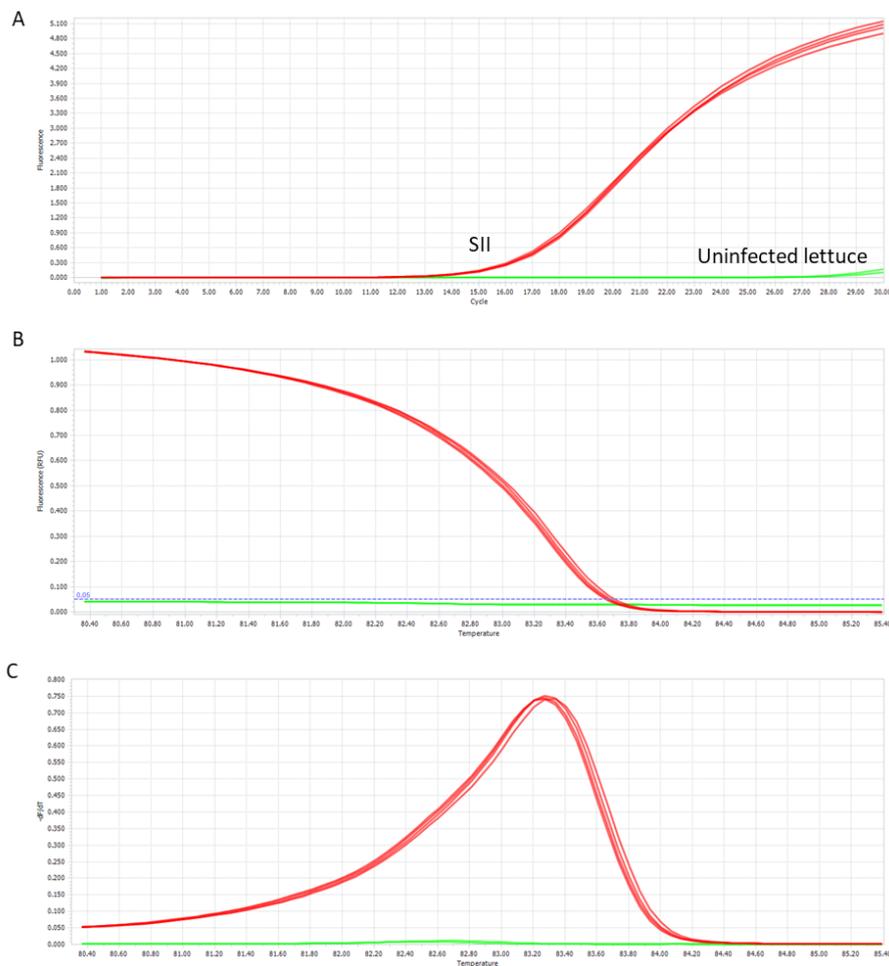
Figure 2.13 Testing the LNYV specificity of the subgroup II Ajithkumar Primers in a singleplex two-step RT-qPCR-HRM analysis with BioRad EvaGreen dye. Subgroup II technical replicates (SII, HV19) are shown in red. Uninfected lettuce technical replicates are shown in light green. A) Amplification curve (Cycle number versus fluorescence). B) Normalised melt curve (Temperature versus fluorescence)C) Normalised HRM melt peaks derived from normalised data shown in B) (Temperature versus -dF/dT). Showing expected subgroup II amplification and unexpected amplification of two uninfected lettuce technical replicates.

Taken together, the results from the above analyses indicate the Ajithkumar Primers are unsuitable for differentiation of the LNYV subgroups in a two-step RT-qPCR-HRM approach. Development of primers specifically for use in two-step RT-qPCR-HRM analysis could provide sensitive differentiation of the two LNYV subgroups. Nevertheless, these analyses provided valuable analysis of the sensitivity of the BioRad Sso Fast EvaGreen dye compared to the QuantaBio's AccuMelt and Invitrogen's SYBR Green fluorescent dyes. For further experiments the BioRad Sso Fast EvaGreen dye was used.

2.3.4 Primer Design for Use in Two-step RT-qPCR-HRM Analysis with EvaGreen Dye

Designing primers specific to each of the two LNYV subgroups proved difficult. There were few regions within the LNYV N gene that allowed subgroup specific binding and that satisfied other primer design criteria, as well as producing PCR products from each subgroup that were sufficiently different in predicted melting temperature to be used for discriminating the LNYV subgroups in HRM analysis. This difficulty was also described with the design of the Ajithkumar Primers (Ajithkumar 2018). Ultimately, three primer sets were settled upon for testing. The primer parameters are displayed in Table 2.2 (Section 2.2.6). Figure 2.14 shows an illustration representing the relative positions on the LNYV N gene of each developed primer set, as well as primers designed previously.
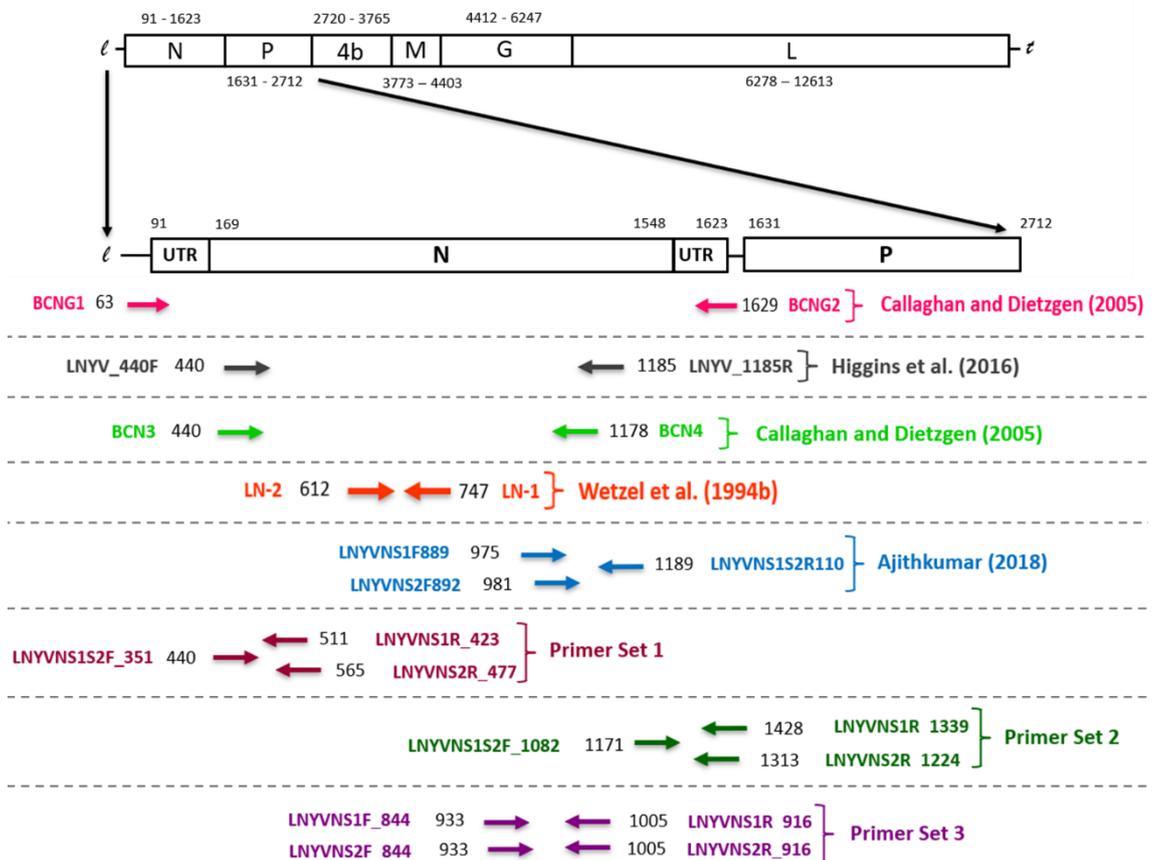
Figure 2.14 Diagram illustrating the positions of the developed primer sets on the LNYV N gene with other primer sets used previous (Figure 2.1). Primer Set 1 (LNYVNS1S2F_351/ LNYVNS1R_423/ LNYVNS2R_477), plum purple, Primer Set 2 (LNYVNS1S2F_1082/ LNYVNS1R_1339/ LNYVNS2R_1224), dark green, and Primer Set 3 (LNYVNS1F_844/ LNYVNS1R_916/ LNYVNS2F_844/ LNYVNS2R_916), purple, were designed in this study for use in two-step RT-qPCR-HRM assays. Ajithkumar Primers and Primer Sets 1, 2 and 3 are named as per the primer positions in the N gene sequence alignment, rather than whole genome position (black numbers). See Table 2.3 for designed primer PCR product sizes, and Figure 2.1 for description of previously used primer sets.

The predicted PCR product properties were also assessed for their suitability for use in the post-PCR HRM analysis. These properties are displayed in Table 2.3. For each primer set, the temperature difference between the subgroup specific PCR products was analysed to ensure their usefulness in subgroup diagnosis by HRM analysis. These are all above the acceptable discrimination difference of 1 °C (Cheng et al. 2013; Osathanunkul et al. 2016), further indicating their suitability. Furthermore, Primer Set 2 shows an estimated temperature difference of 2 °C, which is the desirable temperature for discrimination (Cheng et al. 2013; Osathanunkul et al. 2016).

Table 2.3: Properties of the predicted PCR products for each primer set designed for two-step RT-qPCR-HRM analysis.

Primer name		Subgroup	Expected size	GC%	Predicted T_m (°C)	T_m (°C) difference between subgroups
Primer Set 1	LNYVNS1S2F_351	I	73	46.6	77.8	1.6
	LNYVNS1R_423					
	LNYVNS1S2F_351	II	127	43.3	79.4	
	LNYVNS2R_477					
Primer Set 2	LNYVNS1S2F_1082	I	143	49	79.5	2
	LNYVNS1R_1339					
	LNYVNS1S2F_1082	II	258	44.6	77.5	
	LNYVNS2R_1224					
Primer Set 3	LNYVNS1F_844	I	73	46.6	78.8	1.5
	LNYVNS1R_916					
	LNYVNS2F_844	II	73	56.2	80.3	
	LNYVNS2R_916					

The header "PCR product" spans the columns: Expected size, GC%, Predicted T_m (°C).

2.3.4.1 Testing the Suitability of Primer Set 1 in Two-Step RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

In order to test the suitability of Primer Set 1 for subgroup diagnosis in two-step RT-qPCR-HRM analysis, singleplex (Figure 2.15) and multiplex (Figure 2.16) assays were performed. Primer Set 1 consists of three primers, a non-subgroup specific forward primer (LNYVNS1S2F_351) and two subgroup specific reverse primers (LNYVNS1R_423/ LNYVNS2R_477). Singleplex assays included the common forward primer with either the subgroup I, or subgroup II, specific primer, while the multiplex assay included all three primers. This represents step 3.4 in the experimental design summary (Figure 2.3).

2.3.4.1.1 *Subgroup I and Subgroup II Singleplex Analysis of Primer Set 1*

The singleplex analysis of the subgroup I and subgroup II specific primer pairs of Primer Set 1 are shown in Figure 2.15. For each subgroup specific analysis, amplification of the appropriate subgroup sample was seen; the subgroup I (HV14, blue) technical replicates with the subgroup I primers (Figure 2.15A), and the subgroup II (HV18, red) technical replicates with the subgroup II primers (Figure 2.15B), respectively. This indicates that both subgroup specific primer pairs are specific to their respective LNYV subgroups, and both showed overlap of the amplification profiles during the exponential phase, although subgroup II (HV18, red) showed a much tighter profile. No amplification was seen with the uninfected lettuce, or NTC technical replicates. This indicates there was no non-specific amplification with lettuce, and no contamination in the reactions for both singleplex assays.

The normalised HRM melt peak analysis (Figure 2.15C and Figure 2.15D) showed a single peak with each subgroup primer pair, suggesting one PCR product was amplified for each. However, the subgroup II (HV18, red) melt peak (Figure 2.15D) appeared to show a slight left shoulder leading up to the peak. For both subgroups, variation was seen in the HRM melt peak profiles, although the melt peak temperature ranges were in fact narrow. The subgroup I (HV14, blue; Figure 2.15C) melt peak temperature from the subgroup I primer pair was 78.61 °C to 78.68 °C, while the subgroup II (HV18, red; Figure 2.15D) melt peak temperature from the subgroup II primer pair was 80.87 °C to 80.94 °C, both well within the instrument limit of detection.
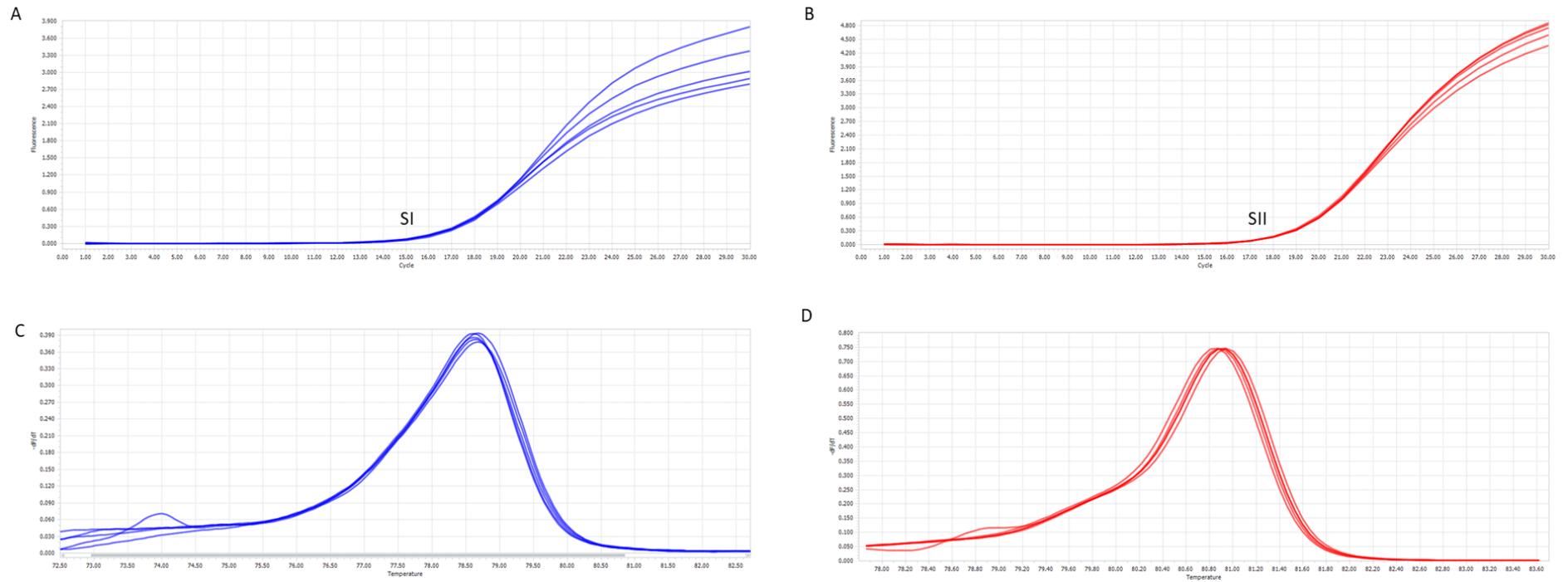
Figure 2.15 Testing the suitability of Primer Set 1 LNYV subgroup specific primers in singleplex assays with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue and subgroup II technical replicates (SII, HV18) are shown in red. A) Subgroup I amplification curve. B) Subgroup II amplification curve. C) Subgroup I normalised HRM melt peaks. D) Subgroup II normalised HRM melt peaks. Showing subgroup specific amplification and one HRM melt peak for each assay. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.4.1.2 *Multiplex Analysis of Primer Set 1*

To confirm the LNYV subgroup specificity of Primer Set 1 when all three primers were combined, a multiplex two-step RT-qPCR-HRM assay was carried out, representing step 3.4 in the experimental design summary (Figure 2.3).

The amplification curve (Figure 2.16A) showed amplification of the subgroup I (HV14, blue), and the subgroup II (HV18, red and orange) technical replicates. Two distinct amplification profiles were observed, one for subgroup I (HV14, blue), and one for subgroup II (HV18, red and orange), although one subgroup II technical replicate showed a differing amplification curve (orange). While equal amounts of RNA were loaded into each reaction for both subgroups, subgroup I amplified earlier. Variation was seen in the amplification profiles for each subgroup, suggesting different amplification efficiencies with greatest variation towards the end of the reactions, as is often observed for PCR (Svec et al. 2015). No amplification of the uninfected lettuce nor the NTC technical replicates was seen.

The normalised HRM melt peak analysis (Figure 2.16B) showed two discrete, non-overlapping melt peaks. The subgroup I (HV14, blue) technical replicates gave a melting temperature range of 78.58 °C to 78.72 °C, while the subgroup II (HV18, red and orange) range was 80.77 °C to 80.90 °C. Both subgroup melt peak temperature ranges were within the instrument's 0.2 °C limit of detection, and the temperature difference between the two peaks was 2.05 °C. A temperature difference of this value is considered appropriate for HRM discrimination (Cheng et al. 2013; Osathanunkul et al. 2016) The subgroup II (HV18) replicate that produced the highest melting temperature is indicated in orange, and corresponds to the orange technical replicate in Figure 2.16A.

To further visualise the replicate variation and profile shape differences between the two LNYV subgroup samples, a difference plot was generated (Figure 2.16C). The subgroup I (HV14, blue) technical replicates were used as the baseline. These samples clustered together close to the baseline, showing very little variation. This is desirable and would be expected in an appropriate baseline sample. A clear different between the subgroup I and subgroup II (HV18, red and orange) curves was seen. The subgroup II (HV18, red and orange) variation that was visible in the HRM melt peak (Figure 2.16B) was more refined in the difference plot, with replicates producing a tight melt curve. The outlying subgroup II replicate (orange) still clustered with the other subgroup II replicates (red), creating a consistent profile shape. Taken together, the difference plot (Figure 2.16C) and the HRM melt peak (Figure 2.16B) results would suggest that the outlying subgroup II product was amplified from an LNYV subgroup II sequence, but had amplified in a dissimilar manner to the other subgroup II replicates.
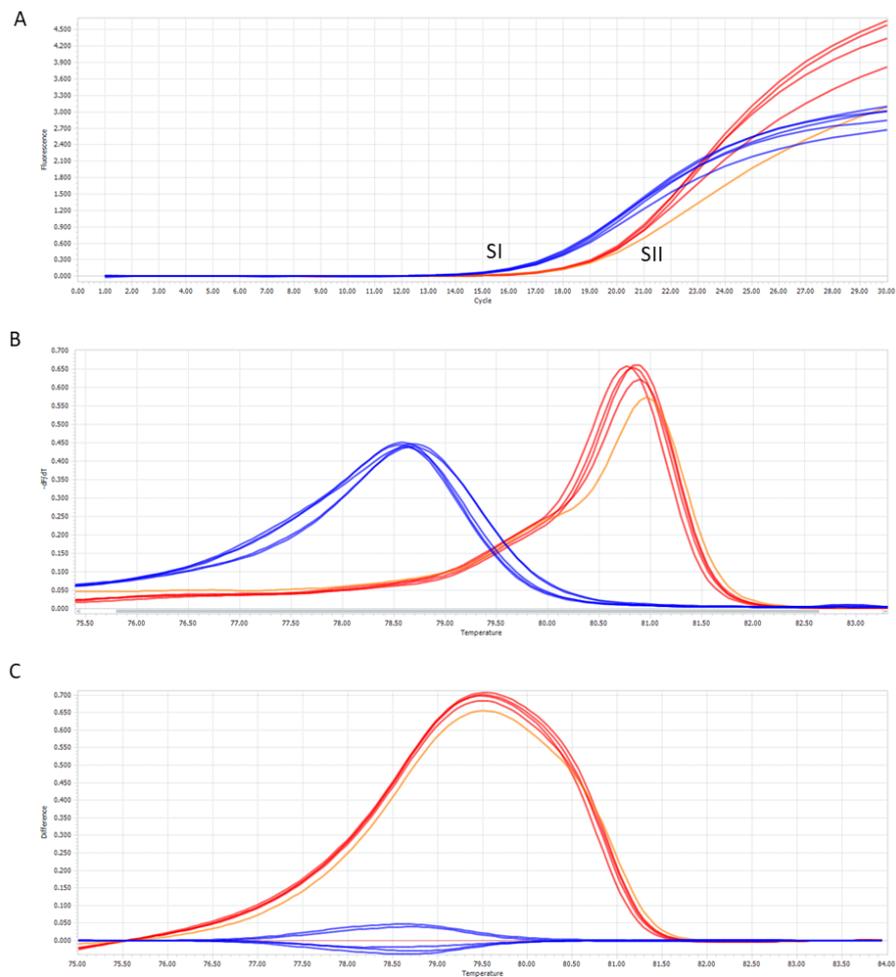
Figure 2.16 Testing the suitability of Primer Set 1 in a multiplex assay with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue and subgroup II technical replicates (SII, HV18) are shown in red and orange. A) Amplification curve (Cycle number versus fluorescence). B) Normalised HRM melt peaks (Temperature versus -dF/dT). C) Difference plot derived from the normalised HRM met peak data shown in B) with the subgroup I technical replicates used as the baseline (Temperature versus difference in -dF/dT). Showing discrete amplification and HRM melt profiles for each subgroup.

The results from both the singleplex (Section 2.3.4.1.1) and multiplex (Figure 2.16) assays indicated that Primer Set 1 did not amplify any non-specific products, and was LNYV subgroup specific. The HRM melt peaks produced from the multiplex assay (Figure 2.16B) indicated this primer combination could distinguish the two LNYV subgroups, and therefore be suitable for quick, cost-effective subgroup diagnosis of samples using a two-step RT-qPCR-HRM approach. For robust analysis of this primer set, testing of untyped samples was the next stage, and is described in Section 2.3.4.5; representing step 3.7 in the experimental design summary (Figure 2.3).

2.3.4.2 Testing the Suitability of Primer Set 2 in Two-Step RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

Singleplex (Figure 2.17) and multiplex (Figure 2.18) analysis with Primer Set 2 was performed in order to test the suitability of the primers for subgroup diagnosis in two-step RT-qPCR-HRM analysis. This represents step 3.5 in the experimental design summary (Figure 2.3). Primer Set 2 consisted of three primers, a non-subgroup specific forward primer (LNYVNS1S2F_1082) and two subgroup specific reverse primers (LNYVNS1R_1339/ LNYVNS2R_1224). Singleplex assays included the common forward primer with either the subgroup I or subgroup II specific primer. The multiplex assay included all three primers.

2.3.4.2.1 *Subgroup I and Subgroup II Singleplex Analysis of Primer Set 2*

The outcomes of the singleplex subgroup specific analysis for Primer Set 2 are shown in Figure 2.17. Both amplification curves (Figure 2.17A and Figure 2.17B) showed amplification of their specific subgroups. However, in the subgroup I specific assay (Figure 2.17A) amplification at cycle 22 was seen in one of the five NTC technical replicates (grey), indicating possible contamination during the qPCR set up. As the remaining four NTC technical replicates were contamination free, it was assumed that this contamination was limited to a single reaction and the assay was not repeated. For both subgroup specific assays, no amplification with uninfected lettuce was seen. The subgroup II (HV18, red) amplification profile (Figure 2.17B) showed tight overlapping of the technical replicates, and there was no contamination in these reactions as indicated by no amplification of the NTC technical replicates.

For the subgroup I (HV14, blue) normalised HRM melt peak (Figure 2.17C) two peaks were seen. One with the subgroup I technical replicates and one from the single NTC technical replicate (grey). The subgroup I (HV14, blue) replicates showed variation, generating a narrow melt peak temperature range of 82.90 °C to 82.97 °C, while the NTC melt peak was at 82.64 °C. There was a temperature difference greater the instrument's 0.2 °C limit of detection between the two peaks, suggesting the contamination had a different sequence from the subgroup I (HV14, blue) sample.

The normalised HRM melt peaks with the subgroup II specific primer pair (Figure 2.17D), showed a single melt peak with a visible left shoulder. Although this shoulder was more defined than the shoulder seen with the Primer Set 1 subgroup II singleplex assay (Section 2.3.4.1.1; Figure 2.15D), it could be indicative of multi-phase melting of a single PCR product, as a second, defined melt peak indicative of the melting of multiple PCR products was not seen (Bester et al. 2012; Downey 2014). Further analysis of the PCR product would be required to

confirm this. The single melt peak generated by the subgroup II (HV18, red) technical replicates showed variation, which produced a narrow temperature range from 82.52 °C to 82.71 °C.
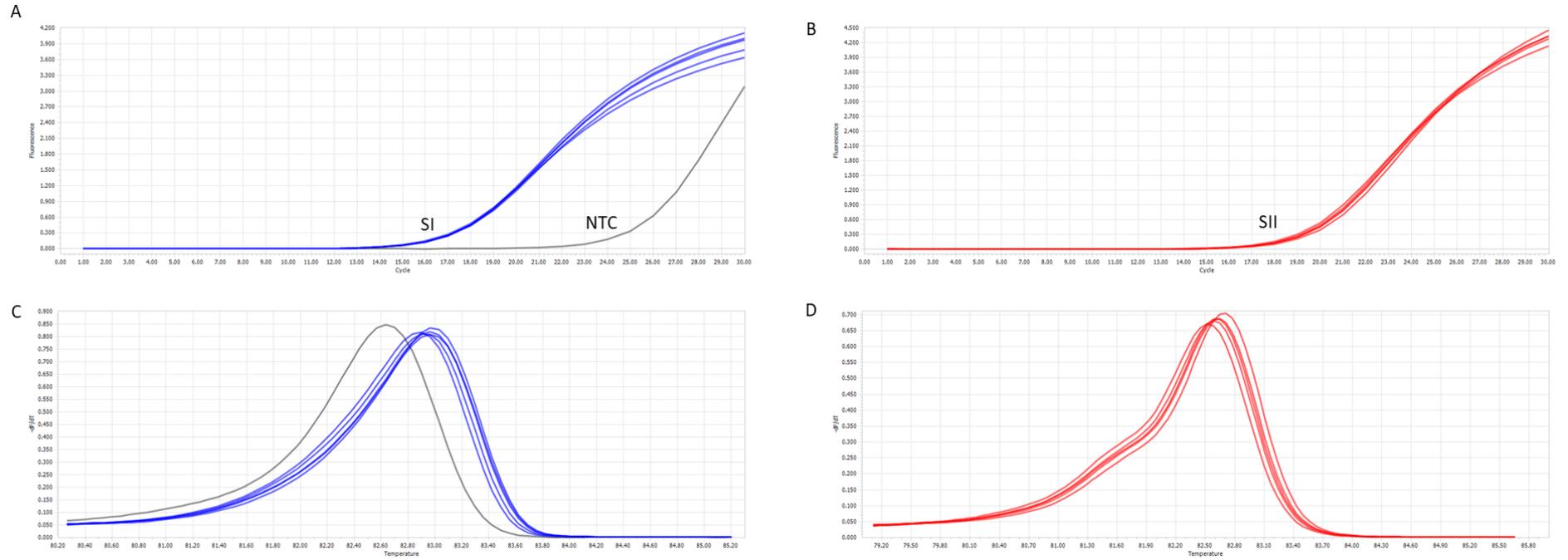
Figure 2.17 Testing the suitability of Primer Set 2 LNYV subgroup specific primers in singleplex assays with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue, the single NTC replicate shown in grey, and subgroup II technical replicates (SII, HV18) are shown in red. A) Subgroup I amplification curve. B) Subgroup II amplification curve. C) Subgroup I normalised HRM melt peaks. D) Subgroup II normalised HRM melt peaks. Showing amplification of both subgroups and possible contamination of one NTC replicate in the subgroup I specific assay. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.4.2.2 *Multiplex Analysis of Primer Set 2*

Singleplex analysis (Section 2.3.4.2.1) of the subgroup specific primer pairs of Primer Set 2 indicated LNYV subgroup specificity. Although there was contamination in one NTC reaction, no other non-specific amplification was seen in either the subgroup I or the subgroup II reaction. However, the melt peak temperatures for both subgroups were similar. Multiplex analysis was carried out as a concluding experiment to test the suitability of Primer Set 2 for LNYV subgroup diagnosis.

The amplification curve (Figure 2.18A) showed amplification of both the subgroup I (HV14, blue), and subgroup II (HV18, red) technical replicates. Tight overlapping amplification curves were seen for each subgroup. Amplification of the subgroup I (HV14, blue) technical replicates occurred slightly earlier than the subgroup II (HV18, red) technical replicates. No amplification was seen with the uninfected lettuce or NTC, indicating there was no contamination or non-specific binding in the multiplex assay.

The normalised HRM melt peak (Figure 2.18B) showed the two subgroups produced melt peaks that were overlapping. Variation in the sample replicates for each subgroup was seen, with subgroup I (HV14, blue) producing a narrower melt peak temperature range of 82.82 °C to 82.89 °C, while for subgroup II (HV18, red) the range was from 82.56 °C to 82.69 °C. The temperature difference between the two peaks was less than the instrument's limit of detection, well below the acceptable temperature difference for discriminating HRM melt peaks (Cheng et al. 2013; Osathanunkul et al. 2016). The subgroup II (HV18, red) melt peak also showed a left shoulder, consistent with the singleplex analysis (Figure 2.17).

A difference plot (Figure 2.18C) was examined to further assess any profile differences between the two subgroups. This emphasised the variation across both subgroup sample replicates. The subgroup I (HV14, blue) technical replicates were used as the baseline, but did not cluster tightly along this baseline as expected. The subgroup II (HV18, red) sample replicates produced melt curves that were spread out, highlighting their variation.
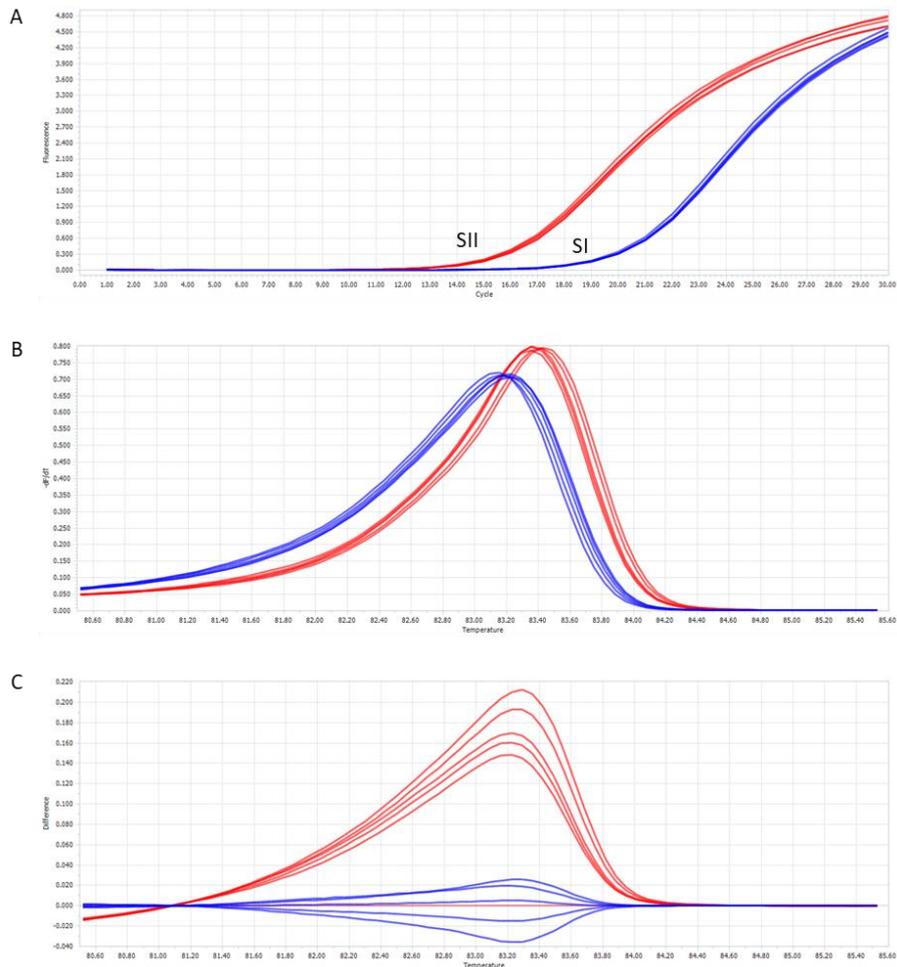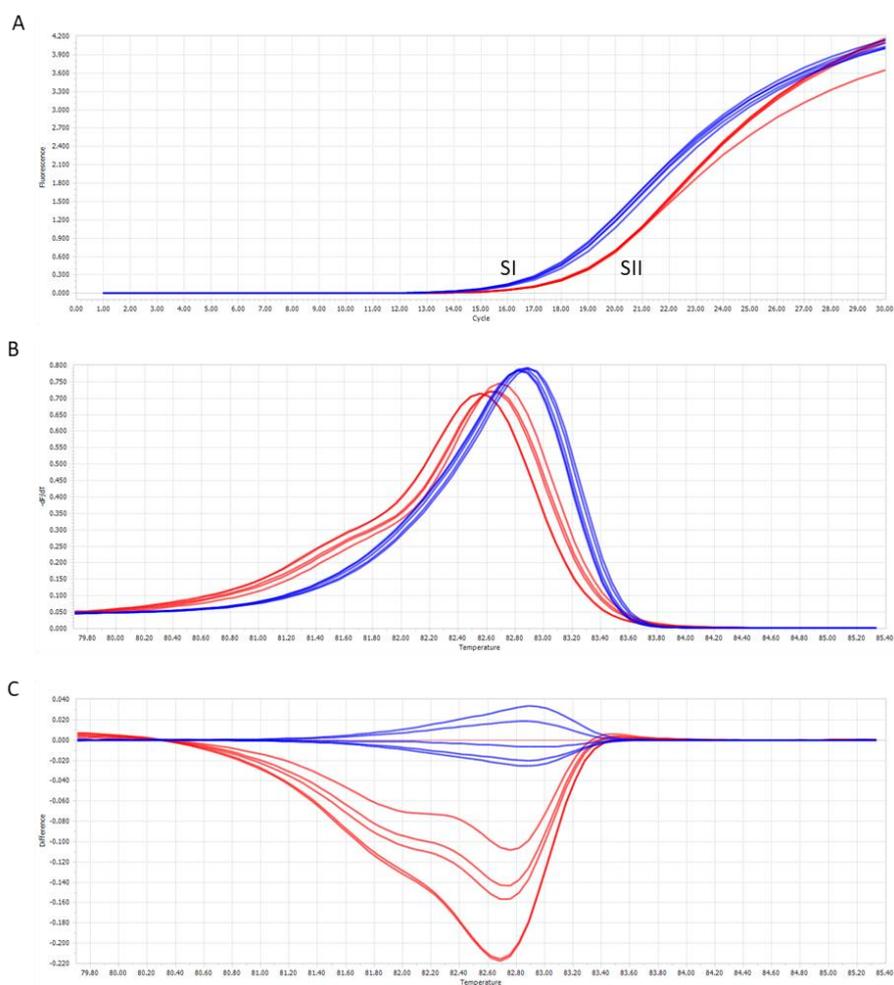
Figure 2.18 Testing the suitability of Primer Set 2 in a multiplex assay with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue and subgroup II technical replicates (SII, HV18) are shown in red. A) Amplification curve (Cycle number versus fluorescence). B) Normalised HRM melt peaks (Temperature versus -dF/dT). C) Difference plot derived from the normalised HRM met peak data shown in B) with the subgroup I technical replicates used as the baseline (Temperature versus difference in -dF/dT). Shows overlapping melt peaks and variation across the technical replicates for each subgroup.

Taken together, the singleplex (Section 2.3.4.2.1) and multiplex analysis of Primer Set 2 indicates these primers are subgroup specific, but are unable to distinguish the two subgroups via HRM analysis. This means Primer Set 2 is unsuitable for LNYV subgroup diagnosis using a two-step RT-qPCR-HRM approach.

2.3.4.3 Testing Suitability of Primer Set 3 in Two-Step RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

Representing step 3.5 in the experimental design summary (Figure 2.3), Primer Set 3 was used in singleplex (Figure 2.19) and multiplex (Figure 2.20) assays to determine its usefulness for

LNYV subgroup diagnosis using two-step RT-qPCR-HRM analysis. Primer Set 3 consisted of four primers, the subgroup I specific forward (LNYVNS1F_844) and reverse (LNYVNS1R_916) primers, and the subgroup II specific forward (LNYVNS2F_844) and reverse (LNYVNS2R_916) primers. Singleplex analysis included each subgroup specific pair, respectively, and multiplex analysis included all four primers.

2.3.4.3.1 *Subgroup I and Subgroup II Singleplex Analysis of Primer Set 3*

The singleplex analysis of the subgroup I and subgroup II specific primer pairs of Primer Set 3 are shown in Figure 2.19. The singleplex subgroup I amplification curve (Figure 2.19A) showed unexpected results. Although tight, overlapping amplification curves were seen with the subgroup I (HV14, blue) technical replicates, amplification was also seen with the subgroup II (HV18, red) technical replicates at cycle 28. The late cycle amplification of the subgroup II (HV18, red) technical replicates suggest low PCR efficiency of these primers with the subgroup II (HV18, red) sample. There was no amplification of the uninfected lettuce nor the NTC technical replicates, indicating there was no contamination in the reactions.

For the subgroup II singleplex analysis, the amplification curve (Figure 2.19B) showed tight, overlapping curves of the subgroup II (HV18, red) technical replicates. There was no amplification of the subgroup I, uninfected lettuce, nor the NTC technical replicates. This indicates subgroup specificity of the subgroup II primer pairs and no assay contamination.

It was discovered in the subgroup I normalised HRM melt peak (Figure 2.19C) that the amplified subgroup I (HV14, blue) and subgroup II (HV18, red) technical replicates produced melt peaks with a similar melting temperature. The subgroup I (HV14, blue) melt peak had a left shoulder, with a narrow melt peak temperature range from 79.93 °C to 80.00 °C. Two of the five subgroup II (HV18, red) technical replicates formed a melt peak, without a shoulder, at 79.87 °C and 80.00 °C. The remaining three subgroup II (HV18, red) sample replicates were normalised to the baseline. The melting temperature for subgroup II samples with Primer Set 3 was predicted to be 1.5 °C higher (Table 2.3) than that for subgroup I. This would suggest the amplification seen was not from contamination with the Primer Set 3 subgroup II specific primers, and could possibly be due to the subgroup I specific primers binding to this particular subgroup II (HV18, red) sequence. Unspecific primer binding has been reported due to nucleotide mismatches at the 3' end of primers (Simsek and Adnan 2000; Lorenz 2012). Further analysis of the primer and subgroup II (HV18, red) sample sequences are required to confirm this. Alternatively, optimising the annealing temperature, or adjusting the qPCR cycle number could reduce the unspecific binding seen with these primers (Lorenz 2012).

A single HRM melt peak was seen with the subgroup II primers in singleplex (Figure 2.19D). This peak consisted of tight, overlapping curves of the subgroup II (HV18, red) technical replicates. This was also reflected in the narrow melt peak temperature range of 83.49 °C to 83.56 °C.

The efficient amplification seen with the subgroup II specific singleplex analysis (Figure 2.19B) and the HRM melt peak temperature (Figure 2.19D) provide further indication that the unspecific binding of the subgroup I primers with the subgroup II (HV18, red) technical replicates (Figure 2.19A and Figure 2.19C) was not due to contamination of the subgroup II specific primers in these reactions.
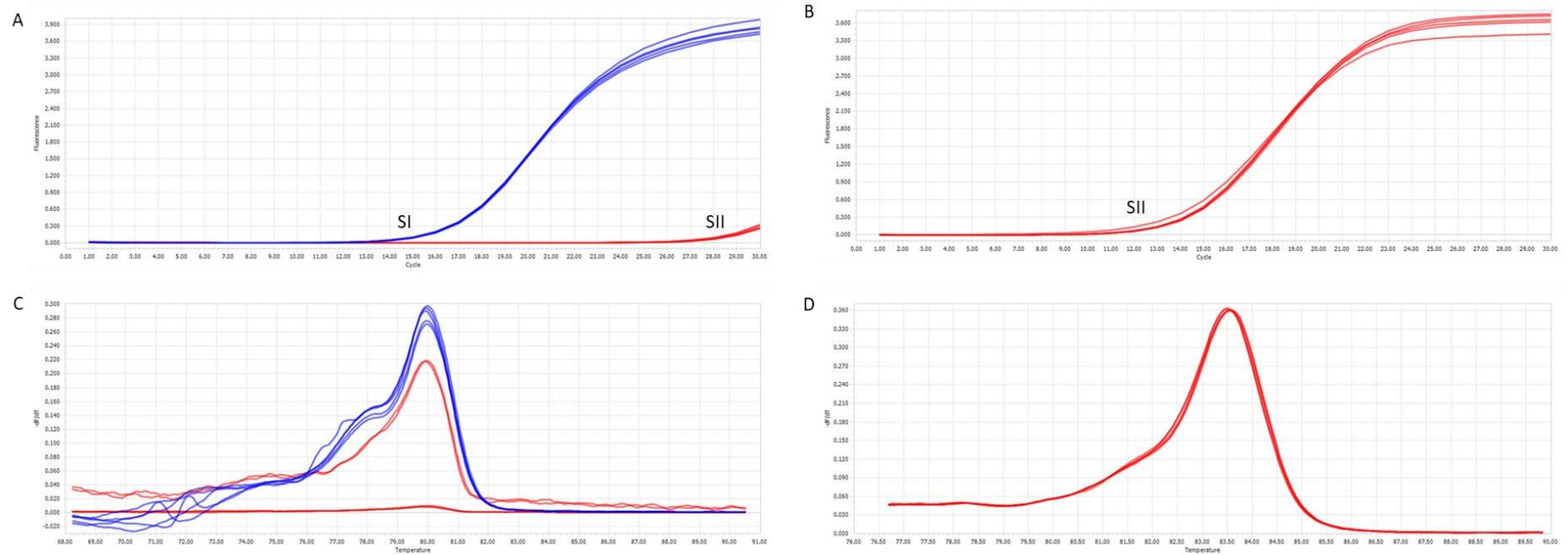
Figure 2.19 Testing the suitability of Primer Set 3 LNYV subgroup specific primers in singleplex assays with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue, and subgroup II technical replicates (SII, HV18) are shown in red. A) Subgroup I amplification curve. B) Subgroup II amplification curve. C) Subgroup I normalised HRM melt peaks. D) Subgroup II normalised HRM melt peaks. Showing possible unspecific amplification of the subgroup I specific primers, while the subgroup II specific primers appear subgroup specific. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT.

2.3.4.3.2 *Multiplex Analysis of Primer Set 3*

In order to assess if the subgroup II specific primers may preferentially and more efficiently bind to subgroup II sequences over the subgroup I specific primers, a multiplex assay with all four Primer Set 3 primers was carried out (Figure 2.20).

The amplification curve (Figure 2.20A) produced two separate amplification profiles. The technical replicates for each subgroup produced overlapping curves in the exponential phase. The subgroup II (HV18, red) sample replicates amplified earlier than the subgroup I (HV14, blue) sample replicates, indicating the subgroup II primer pair was more efficient. No amplification was seen with the uninfected lettuce or the NTC technical replicates.

It was discovered that the subgroup II primers did appeared to bind more efficiently and preferentially to the subgroup II sequence. The normalised HRM melt peak (Figure 2.20B) showed two distinct, non-overlapping melt peaks. Variation was seen for both subgroup sample replicates, producing narrow melt temperature ranges. The subgroup I (HV14, blue) melt peak showed a left shoulder, with temperatures ranging from 79.84 °C to 79.97 °C. The subgroup II (HV18, red) melt peak temperatures ranged from 83.41 °C to 83.55 °C. Both melt peak temperature ranges were within the instrument's 0.2 °C limit of detection, with more than 3 °C separating the two peaks.

The difference plot (Figure 2.20C) reflected the narrow variation across the technical replicates for each subgroup sample, as well as the difference between the two subgroups. The subgroup I (HV14, blue) technical replicates lay along the baseline of the different plot, demonstrating their very little variation. Variation was seen with the subgroup II (HV18, red) sample replicates, however, their overall melt curve shape was consistent, and showed a high degree of difference compared to the subgroup I (HV14, blue) baseline replicates.
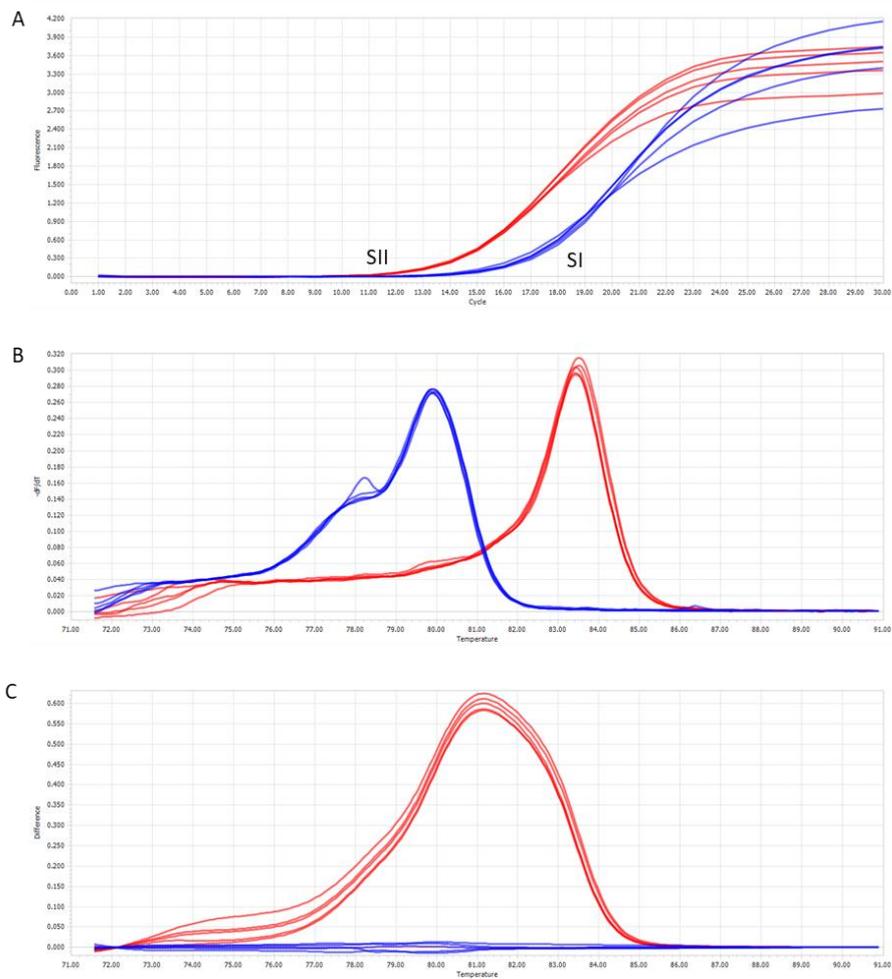
Figure 2.20 Testing the suitability of Primer Set 3 in a multiplex assay with BioRad EvaGreen dye. Subgroup I technical replicates (SI, HV14) are shown in blue and subgroup II technical replicates (SII, HV18) are shown in red. A) Amplification curve (Cycle number versus fluorescence). B) Normalised HRM melt peaks (Temperature versus -dF/dT). C) Difference plot derived from the normalised HRM met peak data shown in B) with subgroup I technical replicates used as the baseline (Temperature versus difference in -dF/dT). Showing subgroup specific amplification and preferential binding of the subgroup II primers to the subgroup II sample.

The singleplex (Section 2.3.4.3.1) and multiplex analysis of Primer Set 3 showed no amplification with uninfected lettuce or NTC samples. Although late, unspecific binding of the subgroup I primers to the subgroup II (HV18, red; Section 2.3.4.3.1) sample was seen. The multiplex analysis indicated the subgroup II primer pair was more efficient, and preferential, than the subgroup I primer pair in binding to the subgroup II (HV18, red) sample, and allowed the two subgroups to be clearly distinguished by HRM analysis. These results indicate, when used in multiplex, that Primer Set 3 is suitable for quick, cost-effective subgroup diagnosis of samples using a two-step RT-qPCR-HRM approach. For robust analysis of this primer set, testing of further untyped samples was the next stage, and is described in Section 2.3.4.5.2, representing step 3.7 in the experimental design summary (Figure 2.3).

2.3.4.4 Comparing the Suitability of Primer Sets 1, 2 and 3 for LNYV Subgroup Diagnosis

In order to determine the most suitable primer set for LNYV subgroup diagnosis using two-step RT-qPCR-HRM analysis, the results from the multiplexing analysis of the three designed primer sets (Sections 2.3.4.1.2, 2.3.4.2.2 and 2.3.4.3.2) were compared together. All three primer sets produced results indicating LNYV subgroup specificity, with no non-specific binding with uninfected lettuce.

For LNYV subgroup diagnosis using HRM analysis, the melt peak temperatures for each subgroup needed to be readily distinct. While a temperature difference of 1 °C is acceptable, a difference of 2 °C is considered appropriate for robust diagnosis (Cheng et al. 2013; Osathanunkul et al. 2016). Figure 2.21A shows the melt peak temperatures for each technical replicate, from each primer set, and the temperature difference between the two subgroups for each primer set. Primer Set 1 and Primer Set 3 produced appropriate temperature differences greater than 2 °C between the two subgroups. While Primer Set 2 showed less than 0.2 °C difference.

Consistency in the melt peak temperatures of the technical replicates is also important. Figure 2.21B shows the mean melt peak temperature of the technical replicates from each primer set, as well as the spread (°C) of the technical replicate peaks. Primer Set 2 showed the smallest spread of melt peak temperatures, however, the subgroup II (HV18) temperature spread was almost double that of subgroup I (HV14) for this primer set. In addition, the subgroup mean temperatures were too similar to distinguish the two subgroups and be appropriate for LNYV subgroup diagnosis. Primer Set 1 and Primer Set 3 showed more consistent temperature spread, with Primer Set 3 producing almost identical temperature spread for both subgroups.

A
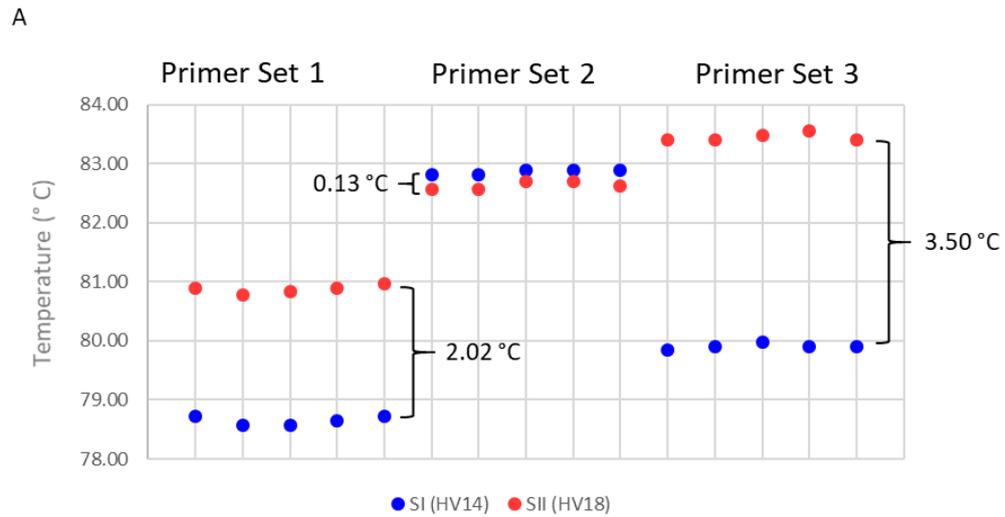
B

	SI (HV14)		SII (HV18)	
	Mean (°C)	Temperature spread (°C)	Mean (°C)	Temperature spread (°C)
Primer Set 1	78.65	0.14	80.87	0.19
Primer Set 2	82.86	0.07	82.62	0.13
Primer Set 3	79.91	0.13	83.45	0.14

Figure 2.21 Comparison of the multiplex HRM melt peak analysis results for Primer Sets 1, 2 and 3 for their suitability for diagnosing the LNYV subgroups. A) Scatter plot showing the normalised HRM melt peak temperatures for each subgroup technical replicate from each primer set. Subgroup I (SI, HV14) is shown in blue and subgroup II (SII, HV18) is shown in red. The temperature (°C) difference between the two subgroups for each primer pair is indicated. B) Table showing the mean melt peak temperatures and the total temperature spread produced by the technical replicates for each subgroup and primer set.

The comparison analysis of the three primer sets confirms that Primer Set 1 and Primer Set 3 are suitable primer sets for LNYV subgroup diagnosis by two-step RT-qPCR-HRM analysis. The next stage was to test these primers with further LNYV samples to conclude robust analysis of these primers and their usefulness in diagnosing the two LNYV subgroups.

2.3.4.5   LNYV Subgroup Diagnosis with Primer Sets 1 and 3 in Two-Step RT-qPCR-HRM Analysis with BioRad EvaGreen Dye

Primer Set 1 and Primer Set 3 were identified as suitable for distinguishing the two LNYV subgroups using two-step RT-qPCR-HRM analysis. In order to complete robust assessment of these primers, they were used to diagnose the subgroup of 18 samples positive for LNYV but had not been subgroup typed. The locations of these samples are illustrated in Figure 2.2. For

these experiments, the positive samples - of known subgroup - were HV14 for subgroup I and HV18 for subgroup II. Three technical replicates for each sample were tested. These experiments represent step 3.7 in the experimental design summary (Figure 2.3).

2.3.4.5.1 *LNYV Subgroup Diagnosis with Primer Set 1*

Figure 2.22A shows the amplification curve for the subgroup diagnosis of LNYV samples with Primer Set 1. Amplification occurred at differing qPCR cycles and was seen for all samples. This suggests there was different amounts of viral template from sample to sample, or different amplification efficiencies for each sample. However, the technical replicates of each sample showed consistent amplification profiles, indicating equivalent amounts of viral template in the reactions and suggests the amplification efficiencies were comparable across the technical replicates of each sample. There was no amplification of the NTC technical replicates, indicating no contamination in the reactions.

Normalisation for HRM melt curve analysis of the samples is based on the melting temperature and the melt curve profile shape (Roche Diagnostics 2012). Thereby, samples, and generally their technical replicates, are grouped together when they generate similar properties (Roche Diagnostics 2012). The samples from this analysis were grouped from A to I, with each group represented by a different colour (Figure 2.22B). Technical replicates for each sample were generally grouped consistently (Figure 2.22B), suggesting that products from each technical replicate had the same melting temperature and profile shape. The samples RPC1 and RPO4 had two technical replicates that grouped in D, with one RPC1 technical replicate in B, and one RPO4 technical replicate in I. These are highlighted in Figure 2.22B for better visualisation here, but were included in group D (dark green) for the rest of the analysis. This would suggest amplification of at least two different sequences from these samples. However, all of these technical replicates clustered with the known subgroup I sample, HV14 (group D, bright blue), as shown by the normalised HRM melt peak (Figure 2.23A).

Sample name	A	A	A	B	B	B	C	C	C	D	D	D	E	E	E	F	F	F	G	G	G	H	H	H	I	I	I
WHG1	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WHG2	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WHG3	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WHG4	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WHG5	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPO1	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPO3	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPE2	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SF3	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPC3	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPE1	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPE3	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SF2	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HV14	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPO4	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-
RPC1	-	-	-	✓	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SF1	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
HV18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-
RPO2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-
RPC2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-	-
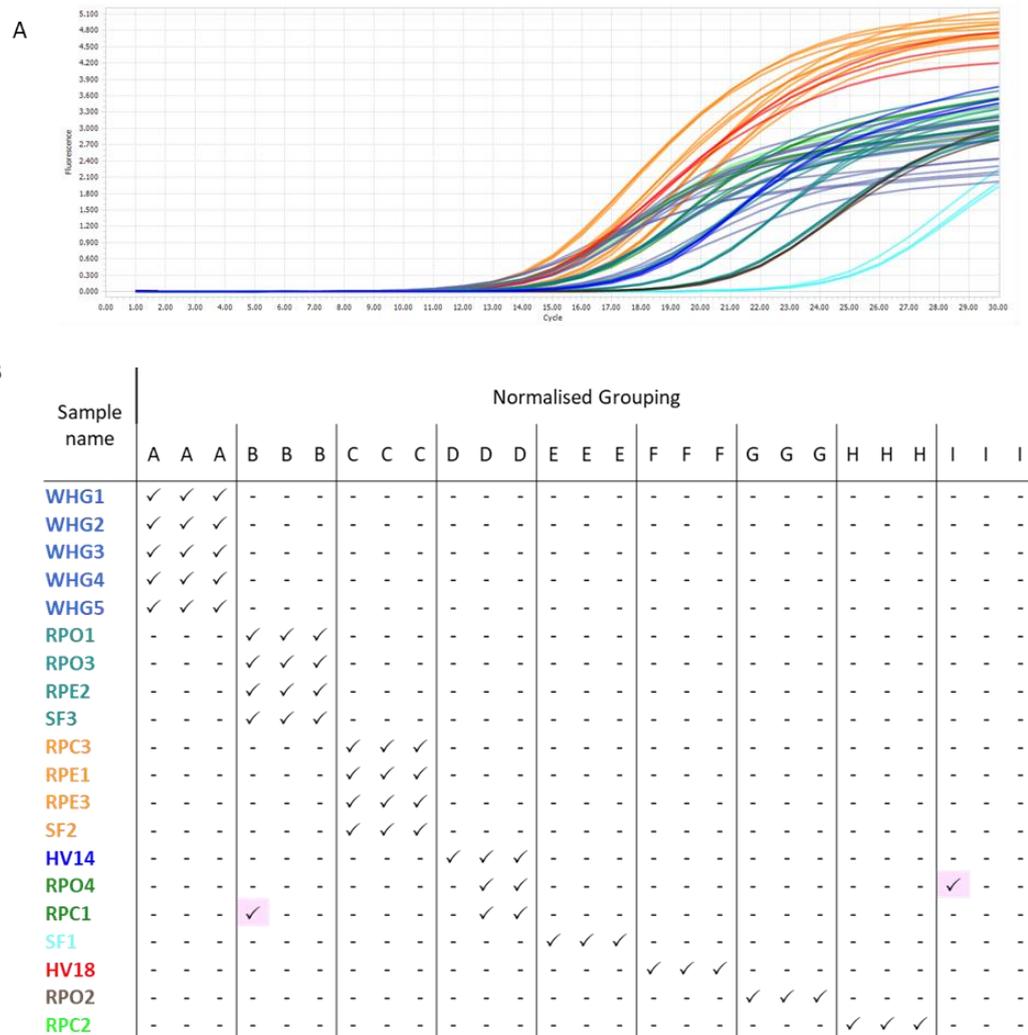
Figure 2.22 LNYV subgroup diagnosis with Primer Set 1. Subgroup I technical replicates (HV14) are shown in bright blue and subgroup II technical replicates (HV18) are shown in red. A) Amplification curve (Cycle number versus fluorescence). Untyped samples are coloured as per their normalised groupings in B). B) Sample technical replicate groupings after normalisation. Samples are coloured by group (A – H), with the exception of HV14 (SI; bright blue). Highlighted technical replicates are those that grouped differently to the other technical replicates for the samples RPO4 and RPC1.

The normalised HRM melt peak (Figure 2.23A) showed that the samples formed two general clusters of samples with discrete melt peak profiles. Sample groups A, B, D, E (Ea), H and G clustered with subgroup I (HV14, group D, bright blue). While sample group C and E (Eb) clustered with the known subgroup II sample, (HV18, group F, red). The samples clustered with subgroup I (HV14, group D, bright blue) produced a general melt peak shape that appeared broader and exhibited less fluorescence when compared to the samples clustered with subgroup II (HV18, group F, red).

With the normalised grouping of samples, the HRM melt peak temperatures were analysed by group rather than each sample individually. Within these groups there was variation of the

melt peaks, although the temperature ranges were mostly narrow (Figure 2.23B). The melt peak temperatures and spread for each normalised group are displayed in Figure 2.23B. One replicate from sample RPC3 (group C, orange) appeared to be an outlier, expanding the temperature variation for group C, and is suggestive of a different sequence for this technical replicate. However, it was grouped consistently with the other RPC3 technical replicates and was within the total subgroup II cluster melt peak temperature range. The subgroup II cluster ranged from 80.10 °C to 80.81 °C. The subgroup II positive control sample (HV18, group F, red) ranged from 80.75 °C to 80.81 °C, and was a consistent temperature with the previous multiplex analysis (Section 2.3.4.1.2). The total subgroup I cluster ranged from 77.95 °C to 78.66 °C, with the subgroup I positive control sample (HV14, group D, bright blue) at 78.66 °C, also consistent with the previous multiplex analysis (Section 2.3.4.1.2). The temperature difference between the highest subgroup I melt peak and the lowest subgroup II melt peak was 1.44 °C, greater than the acceptable difference of 1 °C (Cheng et al. 2013; Osathanunkul et al. 2016).

These results suggest Primer Set 1 is suitable for LNYV subgroup diagnosis. The primers were able to discriminate the LNYV subgroup of samples even with additional LNYV sequences that weren't available during the design of the primers. In fact, sample SF1 (group E, aqua) appeared to contain two sequences. All three technical replicates were normalised into group E, and produced two melt peaks, referred to as Ea, and Eb, respectively. Ea appears to be subgroup I, while Eb appears to be subgroup II (Figure 2.23A and Figure 2.23B). This is the first indication of co-infection by both LNYV subgroups in the same host. To confirm this, the sample was tested with the Ajithkumar Primers and gel electrophoresis and is described in Section 2.3.4.6.

For additional resolution of the normalised sample groupings the difference plot (Figure 2.23C) was examined. The subgroup I (HV14, group D, bright blue) technical replicates were used as the baseline, and these clustered close together, reflecting their single melt peak temperature (78.66 °C). The difference plot showed clear separation of the samples that clustered with subgroup I and those clustered with subgroup II (HV18, group F, red). For SF1 (group E, aqua) the second peak, Eb, was seen above the baseline, while the main melt peak, Ea, was seen below the baseline with the other samples clustered at subgroup I. The outlying RPC3 (group C, orange) replicate was consistent with the other group C and subgroup II samples.
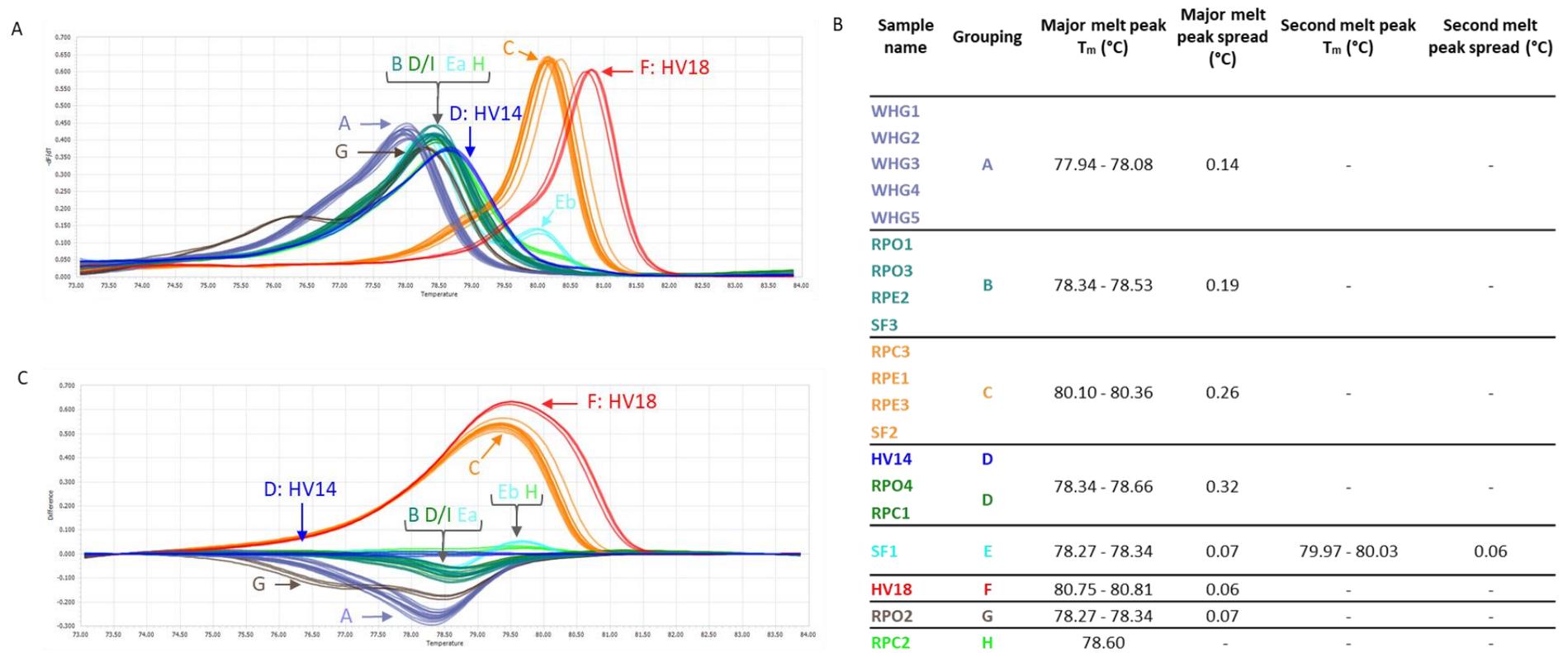
Figure 2.23 HRM analysis of LNYV samples with Primer Set 1. Subgroup I technical replicates (HV14) are shown in bright blue and subgroup II technical replicates (HV18) are shown in red. Untyped samples are coloured as per the normalised groupings in Figure 2.22B. A) Normalised HRM melt peaks (Temperature versus -dF/dT). B) Sample melt peak temperatures by normalised sample group (A – H). C) Difference plot derived from the normalised HRM met peak data shown in A) with subgroup I (HV14) technical replicates used as the baseline (Temperature versus difference in -dF/dT).

Primer Set 1 produced clear subgroup discrimination of previously untyped LNYV samples using two-step RT-qPCR-HRM analysis with BioRad EvaGreen dye. From this data subgroup identification could be carried out, with sample groups A, B, D, E (Ea), G and H indicating LNYV subgroup I, and sample group C indicating LNYV subgroup II. Sample SF1 (group E) indicated the presence of LNYV subgroup II as well as LNYV subgroup I, suggesting possible co-infection of the two subgroups in this host sample. Further confirmation with the Ajithkumar Primers and gel electrophoresis, was carried out and is described in Section 2.3.4.6.

2.3.4.5.2 *LNYV Subgroup Diagnosis with Primer Set 3*

Primer Set 3 also appeared suitable, in multiplex two-step RT-qPCR-HRM analysis, for differentiating between the two LNYV subgroups (Section 2.3.4.4). The 18 previously untyped LNYV samples were also diagnosed with Primer Set 3 (Section 2.3.4.5.2).

The amplification curve (Figure 2.24A) showed consistent amplification profiles for sample technical replicates, but these differed from sample to sample. This suggests the viral template amount across the technical replicates was equivalent, but varied between samples, or each sample showed different amplification efficiency with these primers. This was also seen with the amplification profiles produced with Primer Set 1 (Figure 2.22A). No amplification with the NTC technical replicates was seen, indicating there was no contamination in the reactions.

In contrast to the analysis with Primer Set 1 (Figure 2.23B), samples and their technical replicates did not group consistently, although the normalisation sensitivity settings remained the same. This suggests there were subtle differences in the melt peak temperatures or in the produced profile shapes with Primer Set 3. The sample groupings and representing colours from the Primer Set 1 analysis were used with the Primer Set 3 analysis for consistency and are shown in Figure 2.24B. The positions of samples produced in the normalised HRM melt peak analysis (Figure 2.25A) also supported keeping this consistent.
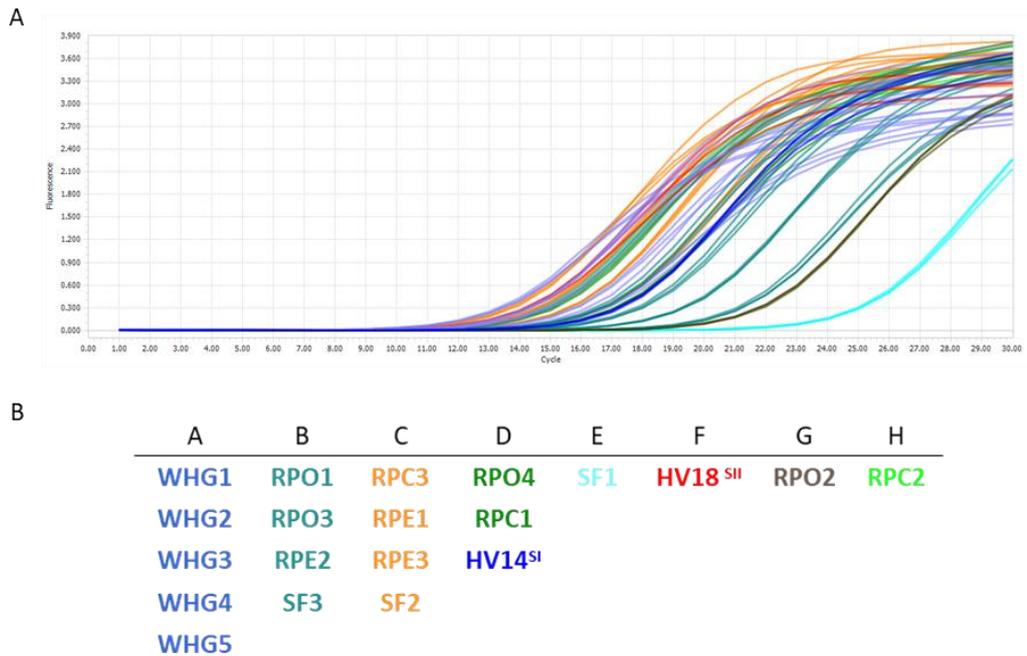
Figure 2.24 LNYV subgroup diagnosis with Primer Set 3. Subgroup I technical replicates (HV14) are shown in bright blue with superscript 'SI' and subgroup II technical replicates (HV18) are shown in red with superscript 'SII'. Untyped samples are coloured as per the normalised groupings in Figure 2.15B. A) Amplification curve (Cycle number versus fluorescence). B) Sample normalised groupings as per Section 2.3.4.5.1; Figure 2.15B.

The normalised HRM melt peaks (Figure 2.25A) showed two discrete sample clusters. Samples that clustered with subgroup I (HV14, group D, bright blue) showed a distinctive left shoulder that was also seen in the Primer Set 3 multiplex testing (Section 2.3.4.3.2). The samples that clustered here; WHG1-5, RPO1-4, RPC1, RPC2, RPE2, SF1 and SF3, are consistent with those identified with the Primer Set 1 diagnosis (Section 2.3.4.5.1), and represent the normalised groups A, B, D, E, G and H (Figure 2.24B). Further, samples RPC2, SF1, and RPC1 produced a second melt peak that was in-line with the subgroup II (HV18, group F, red) cluster, although these peaks generated much lower fluorescence.

In contrast to the Primer Set 1 analysis, the subgroup I sample cluster produced similar fluorescence to the subgroup II sample cluster. The other samples that clustered with subgroup II (HV18, group F, red) were RPC3, RPE1, RPE3, and SF2, and were consistent with the group C samples that were identified as subgroup II in the Primer Set 1 diagnosis (Section 2.3.4.5.1). Sample SF2 (group C, orange) produced a different profile shape with a visible left shoulder.

The samples clustered with subgroup I (HV14, group D, bright blue) that produced a right shoulder or a second melt peak were isolated for better resolution and closer analysis (Figure

2.25C). Clear second peaks were seen for RPC2 (group Hb, light green), SF1 (group Eb, aqua), and RPC1 (group Db, dark green). The subgroup I sample (HV14, group D, bright blue), showed a peak shape but this was very close to the baseline fluorescence. Sample SF3 (group B, teal) was included as it showed a shoulder in this area, but the fluorescence was even lower than that observed for HV14 (subgroup I, group D, bright blue).

The difference plot (Figure 2.25D) was examined to further assess the profile differences between the subgroup I and subgroup II clusters, and in particular those samples that showed a second melt peak. The subgroup I (HV14, group D, bright blue) technical replicates were used as the baseline. Although all samples produced curves above the baseline, there were discrete curves seen for the samples clustered with subgroup I or subgroup II. This indicates clear differences in their HRM profile shapes. The samples that produced a second peak, RPC2 (group Hb, light green), SF1 (group Eb, aqua), and RPC1 (group Db, dark green), showed a different profile compared to the samples that produced only one peak. Sample SF3 (group B, teal), which showed a shoulder rather than a second peak, produced a profile that was consistent with the single peak subgroup I samples so was considered as such for any further analyses. (subgroup I, HV24, bright blue), there was a clear difference between the subgroup I and the subgroup II samples.

The HRM melt peak temperatures (Figure 2.25B) were looked at closely to determine if these had influenced the inconsistent groupings of technical replicates after normalisation. Because of this, the melt peak temperatures were calculated per sample across its technical replicates, rather than per normalised grouping. The complete subgroup I cluster ranged from 80.35 °C to 80.55 °C. The subgroup I positive control sample (HV14, group D, bright blue) was slightly higher than the multiplex analysis (Section 2.3.4.3.2) range of 79.84°C to 79.97 °C, with a melt peak temperature of 80.03 °C. While the subgroup II cluster HRM melt peaks ranged from 83.48 °C to 83.67 °C (Figure 2.25B), and the subgroup II positive control sample (HV18, group F, red) was at 83.55 °C, consistent with the multiplex analysis (Section 2.3.4.3.2). The temperature difference between the two subgroup clusters was 2.94 °C, greater than the desirable difference of 2 °C (Cheng et al. 2013; Osathanunkul et al. 2016). This suggests these primers are well suited to discriminate the two LNYV subgroups using HRM analysis. In addition, with this primer set, second melt peaks were identified for three samples, compared to just one sample using Primer Set 1 (Section 2.3.4.5.1).

For Primer Set 3, 8 of the 20 samples showed no variation in the melt peak temperatures across their technical replicates at all. The temperature variation of the remaining 12 samples, and the three second melt peaks, were all below the instrument's 0.2 °C limit of detection.

These results indicate the primers amplified consistently for all technical replicates, suggesting stable amplification efficiencies or equivalent viral template in each technical replicate. The narrow melt peak temperature variation seen indicates the inconsistent sample groupings after normalisation was due to subtle differences in the profile shape of the technical replicates, rather than differences in melt peak temperatures. This suggests this primer set allows for subtle sequence differences to be amplified, without altering its ability to diagnose the two LNYV subgroups.
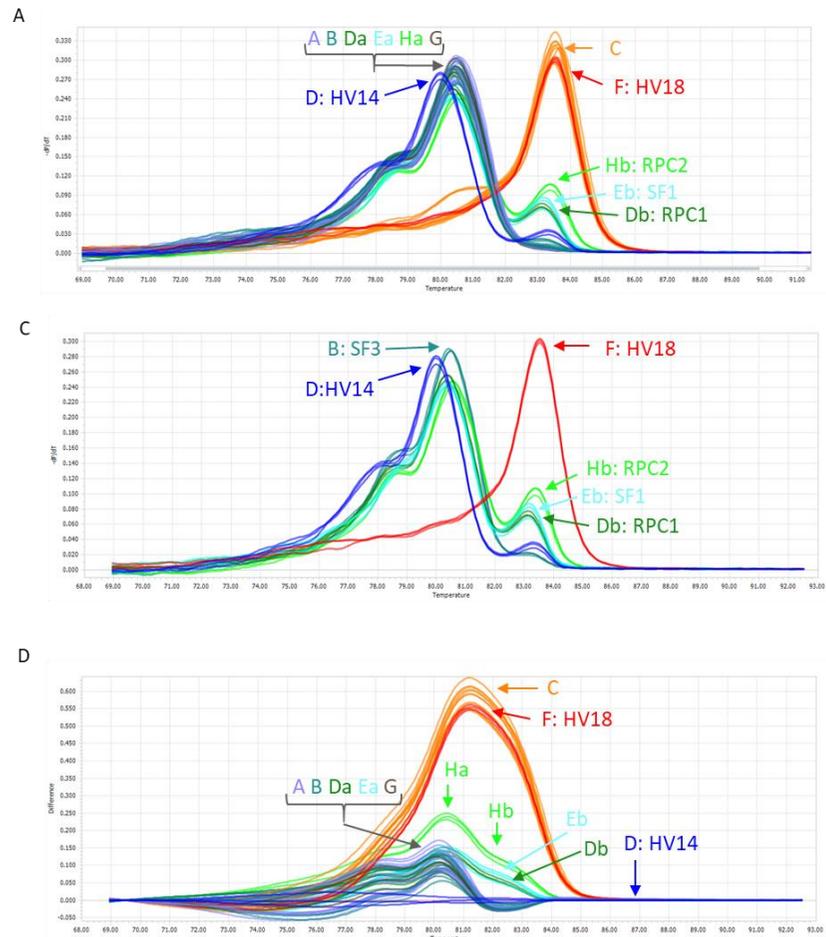
Sample Name	Major melt peak Tm (°C)	Major melt peak spread (°C)	Second melt peak Tm (°C)	Second melt peak spread (°C)
HV14	80.03	-	-	-
WHG1	80.48 - 80.55	0.07	-	-
WHG2	80.42 - 80.49	0.07	-	-
WHG3	80.42 - 80.55	0.13	-	-
WHG4	80.49 - 80.55	0.06	-	-
WHG5	80.42 - 80.55	0.13	-	-
RPO1	80.35	-	-	-
RPO2	80.42 - 80.48	0.06	-	-
RPO3	80.48	-	-	-
RPO4	80.42 - 80.48	0.06	-	-
RPC1	80.35	-	83.09 - 83.16	0.07
RPC2	80.55	-	83.36 - 83.42	0.06
RPC3	83.61 - 83.67	0.06	-	-
HV18	83.55	-	-	-
RPE1	83.55	-	-	-
RPE2	80.48 - 80.55	0.07	-	-
RPE3	83.55	-	-	-
SF1	80.35 - 80.42	0.07	83.15 - 83.22	0.07
SF2	83.48 - 83.55	0.07	-	-
SF3	80.42 - 80.48	0.06	-	-

Figure 2.25 HRM analysis of LNYV samples with Primer Set 3. Subgroup I technical replicates (SI, HV14) are shown in bright blue and subgroup II technical replicates (SII, HV18) are shown in red. Untyped samples are coloured as per Figure 2.24B. A) Normalised HRM melt peaks. B) Sample melt peak temperatures. C) Normalised HRM melt peaks with selected samples. D) Difference plot derived from the normalised HRM melt peak data shown in A) with subgroup I (HV14) technical replicates used as the baseline. A) and C) Temperature versus -dF/dT. D) Temperature versus difference in -dF/dT.

Both Primer Set 1 and Primer Set 3 proved suitable for discriminating the two LNYV subgroups. However, Primer Set 3 showed greater sensitivity when compared with Primer Set 1. Primer Set 3 generated a greater temperature difference between the subgroup I and subgroup II samples, as well as identified three samples showing co-infection of the two subgroups, compared to just one with Primer Set 1. Although the sample replicates did not group consistently after normalisation of the data with Primer Set 3, the small amount of variation in the melt peak temperatures suggests this is due to subtle differences in profile shape, suggesting the region this primer set amplifies allows differences in the sequences to be identified, but does not affect its ability to diagnose between the two subgroups.

2.3.4.6 Confirmation of Co-Infection of LNYV Subgroups I and II Among Samples

In order to confirm the apparent co-infection of samples with the two LNYV subgroups, additional singleplex analysis was carried out. Using the Ajithkumar Primers, two-step RT-qPCR-HRM analysis followed by gel electrophoresis of the three samples identified as co-infected was performed (Figure 2.26). This analysis represents step 3.8 in the experimental design summary (Figure 2.3).

The subgroup I specific amplification curve (Figure 2.26A) showed amplification of all samples, including the subgroup II (HV18, red) technical replicates. Several amplification curves were seen, with amplification occurring at different cycles for each sample. This suggests varied amounts of viral template or different PCR efficiencies from sample to sample, although, amplification of sample technical replicates was consistent. Unexpectedly, late amplification, at cycle 28, was seen with the subgroup II (HV18, red) technical replicates. This could suggest these primers have low qPCR efficiency with this particular subgroup II sample. No amplification was seen with the NTC technical replicates, indicating no contamination in the reactions.

On the other hand, the subgroup II specific amplification curve (Figure 2.26B) showed no amplification of the subgroup I (HV14, bright blue) sample. While the subgroup II (HV18, red) technical replicates amplified consistently, as expected. Consistent late amplification was also seen with RPC2 (light green), RPC1 (dark green), and SF1 (aqua), technical replicates. This assay also showed no contamination in the NTC technical replicates.

The normalised HRM melt peaks (Figure 2.26C and Figure 2.26D) were examined to identify if sufficient amplification had occurred to generate a peak, as it was discovered the Ajithkumar Primers were unable to discriminate the two LNYV subgroups by HRM melt peak temperature

(Sections 2.3.3.5 and 2.3.3.6). The subgroup I specific HRM melt peak (Figure 2.26C) generated peaks for RPC1 (dark green), RPC2 (light green), SF1 (aqua) and the subgroup I sample (HV14, bright blue). The subgroup II (HV18, red) produced only a hump shape near the baseline, most likely due to the late amplification and low fluorescence generated.

The subgroup II specific HRM melt peak (Figure 2.26D) showed peaks for RPC1 (dark green), RPC2 (light green), and the subgroup II (HV18, red) technical replicates. Sample SF1 (aqua) produced only a hump shape near the baseline, most likely due to its late amplification. The PCR products for each technical replicate were assessed by gel electrophoresis due to the HRM unsuitability of the Ajithkumar Primers (Sections 2.3.3.5 and 2.3.3.6).

Gel electrophoresis showed bands of the expected 200 bp size for all samples, including the subgroup II (HV18, red) technical replicates, with the subgroup I specific primers (Figure 2.26E). Bright PCR product bands at were seen with RPC1 (dark green), RPC2 (light green), SF1 (aqua), and the subgroup I (HV14, bright blue) samples. This indicates these samples were infected with LNYV subgroup I. Unexpectedly a faint band was also seen with subgroup II (HV18, red) technical replicates, even though amplification occurred late and only a hump was seen on the HRM melt peak analysis (Figure 2.26C). This suggests the subgroup I specific primers are binding to this particular LNYV subgroup II sample when used in two-step RT-qPCR assays. This could be due to the increased efficiency of qPCR (Gachon et al. 2004), or nucleotide mismatches at the primer 3' end, which have been reported to produce unspecific primer binding (Simsek and Adnan 2000; Lorenz 2012). This was also seen with the subgroup I primer pair from Primer Set 3 (Section 2.3.4.3.1). Further analysis of the primer sequences against the HV18 sample sequence would be a first step towards assessing if unspecific binding had occurred, and is described in Section 2.3.4.7, representing experimental summary design (Figure 2.3) step 4.0. Gel electrophoresis of the NTC technical replicates confirmed no contamination in the reactions.

Gel electrophoresis of the subgroup II specific primers (Figure 2.26F) showed the subgroup II (HV18, red) technical replicates produced the brightest band of the expected size. This was anticipated, and reflects the high qPCR efficiency or the high amount of viral template in this sample, as was suggested by the amplification curve (Figure 2.26B). Bands of the expected 200 bp size were seen for RPC1 (dark green), RPC2 (light green), and also the SF1 (aqua) technical replicates that showed late amplification (Figure 2.26B) and a low hump on the HRM melt peak (Figure 2.26D). These results indicate these samples were also infected with LNYV subgroup II. No bands were seen for the subgroup I (HV14, bright blue) or the NTC technical replicates.

Figure 2.26 Confirming LNYV subgroup co-infection using Ajithkumar Primers, followed by gel electrophoresis. Technical replicates for subgroup I (SI; HV14) are shown in bright blue, subgroup II (SII; HV18) shown in red, RPC1 shown in dark green, RPC2 shown in light green, and SF1 shown in aqua. A) Amplification curve with the subgroup I primer pair. B) Amplification curve with the subgroup II primer pair. C) Normalised HRM melt peaks with the subgroup I primer pair. D) Normalised HRM melt peaks with the subgroup II primer pair. E) Example gel electrophoresis showing one technical replicate for each sample with the subgroup I primer pair. F) Example gel electrophoresis showing one technical replicate for each sample with the subgroup II primer pair. A) and B) Cycle number versus fluorescence. C) and D) Temperature versus -dF/dT. E) SI + represents subgroup I positive control. F) SII + represents subgroup II positive control.

Three lettuce samples (RPC1, RPC2 and SF1) have been identified as co-infected with both LNYV subgroups. This was originally diagnosed by primers developed specifically for two-step RT-qPCR-HRM analysis (Section 2.3.4.5), but was then confirmed using subgroup specific primers designed for use in end point RT-PCR-RFLP (Ajithkumar 2018). Gel electrophoresis of the singleplex analysis with the Ajithkumar Primers designed for end point RT-PCR analysis, showed bands of the correct size. This was also seen for technical replicates that produced late qPCR amplification, and only humps along the baseline in the normalised HRM melt peak analysis (Figure 2.26). To date, this is the first instance of co-infection of the LNYV subgroups in the same host to be reported.

2.3.4.7   Assessment of Primer Sequences on the LNYV N Gene

It was discovered that the subgroup II sample, HV18, showed amplification with the subgroup I specific primers of the Ajithkumar Primers (Section 2.3.4.6) and the designed Primer Set 3 (Section 2.3.4.3.1). This was unexpected, as in both cases the primers were designed to be subgroup I specific. However, the HV18 N gene sequence was not available at the time of primer design, and unspecific primer binding due to nucleotide mismatching at the 3' end of the primers has been reported (Simsek and Adnan 2000; Lorenz 2012). The analysis carried out in Chapter 3 has provided N gene sequences of the subgroup II sample, HV18, as well as the three samples identified as co-infected; RPC1, RPC2 and SF1. These N gene sequences can be used to assess the primer binding of the developed Primer Set 3, and the subgroup specific forward primers of the Ajithkumar Primers (Figure 2.27). This analysis represents experimental design summary (Figure 2.3) step 4.0.

Unspecific binding of primers can occur due to nucleotide mismatching at the 3' end of the primers (Simsek and Adnan 2000; Lorenz 2012). Assessment of the subgroup I specific primers (Figure 2.27A) for both Primer Set 3 (purple arrow) and the Ajithkumar Primers (blue arrow) showed the required six nucleotide differences correlating to subgroup specificity were maintained. This was assessed against the subgroup II (HV18) N gene sequence The subgroup I forward primer for Primer Set 3 (purple arrow) contained three degenerate primers (indicated by the two bases separated by /), and these also did not correlate to the subgroup II (HV18) N gene sequence. However, it has been reported that a G to T mismatch was as efficient as the original primer (Simsek and Adnan 2000). This type of nucleotide mismatch could occur with the first degenerate base, representing T/C, of Primer Set 3 and the G nucleotide of the subgroup II (HV18) sequence. However, this was not observed with the Ajithkumar Primers subgroup I forward primer.

The three samples identified as co-infected were assessed with the subgroup I specific primers (Figure 2.27A) and the (Figure 2.27B). The N gene sequences of these samples correlated with the subgroup I specific primers only. Within the Primer Set 3 forward primer binding sites, these samples showed only one nucleotide that was similar to the subgroup II (HV18) N gene sequence compared to the subgroup I (HV14) N gene sequence (fourth nucleotide within the primer binding site). Within the Ajithkumar Primer binding sites, these samples correlated with the subgroup I (HV14) N gene sequence. RPC2 and RPC1 showed one exception, where one nucleotide (seventh nucleotide within the subgroup I primer site, and fourth nucleotide within the subgroup II primer site) correlated with neither subgroup I (HV14) or subgroup II (HV18) N gene sequences.

A

Ajithkumar Primers SI forward primer
ACC TGA AGT CTT ATC CAC ATG GAC TTT A

Primer Set 3 SI forward primer
T/CAC G/TGA AG AGC CTT AGT G/ATC AT

Primer Set 3 SI reverse primer
ACC TGA AGT CTT ATC CAC ATG GTC TTT A

Identity
Subgroup II (HV18)
Subgroup I (HV14)
SF1
RPC2
RPC1

B

Ajithkumar Primers SII forward primer
C GAA GTG TTA TCG ACA TGG ACA CTG

Primer Set 3 SII forward primer
GAC AGA GGC AGC TTT GGT GTC AC

Primer Set 3 SII reverse primer
C GAA GTG TTA TCG ACA TGG ACA CTG

Identity
Subgroup II (HV18)
Subgroup I (HV14)
SF1
RPC2
RPC1

Figure 2.27 Sequence alignment of Primer Set 3 (purple arrows) and the forward LNYV subgroup specific primers of the Ajithkumar Primers (blue arrows) to select LNYV isolates to assess primer binding sites. The reverse primer for the Ajithkumar Primers is not shown as it was non-specific and binds to both subgroups (Ajithkumar 2018). Primers and sample sequences shown in 3' to 5' direction. Primer Set 3 subgroup I forward primer contains three degenerate primers indicated by the two bases separated by /. Primer Set 3 reverse primers shown as the reverse complement sequence to aid comparison. Identity bar shown to highlight nucleotide differences across the sequences (yellow colour). A) Subgroup I (SI) primer binding sites. B) Subgroup II (SII) primer binding sites.

This data suggests the N gene sequence amplified from the three samples identified as co-infected reflects LNYV subgroup I only. These samples were amplified by one-step RT-PCR (Chapter 3), and perhaps this was the most abundant sequence in the samples, as indicated by the major HRM melt peaks (Sections 2.3.4.5.1 and 2.3.4.5.2). The two-step RT-qPCR analysis performed in this Chapter appears to have amplified other sequences available in each sample's total RNA. This is shown by amplification with the subgroup II primers of both Primer Set 3 and the Ajithkumar Primers (Sections 2.3.4.5.2 and 2.3.4.6), while no correlation was seen with the subgroup II primer binding sites and the N gene sequences of the samples.

2.4 Discussion

This Chapter describes the development of an LNYV subgroup specific diagnostic assay using two-step RT-qPCR-HRM analysis. As suggested by Ajithkumar (2018), the primers originally developed for RT-PCR-RFLP encompass several features that indicated their potential suitability for use in RT-qPCR-HRM assays. It was discovered that the one-step RT-qPCR-HRM analysis using the Invitrogen SYBR Green dye with the Ajithkumar Primers (Section 2.3.2.1) was unsuitable for subgroup discrimination. The results required further confirmation as amplification profiles and the normalised HRM melt peaks were inconsistent across the technical replicates for each sample, even though the same amount of total RNA was used. This suggested that the PCR products of the technical replicates may have different sequences that may reflect that LNYV exists as a quasi-species (expanded on below). However, gel electrophoresis confirmed the PCR products for all technical replicates were of the expected size.

The LightCycler 96 System includes a specific HRM module, and the instrument is also combatable with the SYBR Green fluorescent wavelength (Roche Diagnostics 2012). This indicates fluorescent measurement was not the source of the inconsistent HRM melt profiles, and suggests the SYBR Green dye reagent itself could be the source of the unexpected variation seen. Although commonly used in qPCR experiments, several studies have indicated the unsuitability of SYBR Green dye in qPCR, and melt curve or HRM analysis (Varga and James 2006; Gudnason et al. 2007; Eischeid 2011; Taylor et al. 2011; Radvanszky et al. 2015; Pereira et al. 2018). In fact, a study in 2015 (Radvanszky et al.) which aimed to compare 12 green fluorescent dyes for their suitability in HRM applications, eliminated SYBR Green during the first round of the study. This was due to its poor performance regarding fluorescence after binding to dsDNA, its PCR amplification inhibition, and its effect on PCR product melting temperatures (Radvanszky et al. 2015).

The poor performance of SYBR Green dye in qPCR-HRM assays has been attributed to its non-saturating nature and subsequent translocation (Figure 2.28; Varga and James 2006; Li et al. 2010; Eischeid 2011; Radvanszky et al. 2015; Pereira et al. 2018). Dyes designed specifically for use in qPCR-HRM assays; such as the QuantaBio AccuMelt, containing SYTO 9, and BioRad EvaGreen dyes, are saturating DNA binding dyes (Biotium 2009; Quanta Bioscience 2010; Bio-Rad 2014).The saturating nature of these dyes means, as the dsDNA is melted, the dye is unable to translocate to other double-stranded regions, causing the fluorophore to become unexcited, and therefore, the fluorescence decreases (Figure 2.28; Taylor et al. 2011; Pereira et al. 2018).
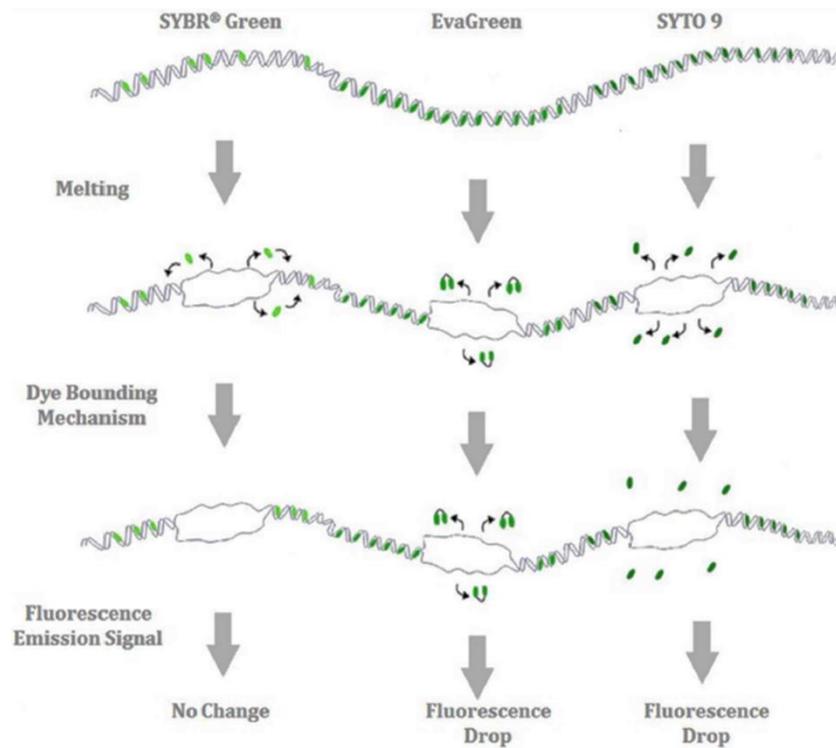
Figure 2.28 Diagram illustrating saturating and non-saturating nature of fluorescent DNA binding dyes. Common intercalating dyes used for HRM analysis include SYBR Green, EvaGreen and SYTO 9. The non-saturating nature of SYBR Green dye allows the dye to translocate to other un-melted, double-stranded regions of the PCR product, resulting in less or no change in fluorescent signal. This is shown by the black arrows showing the movement of dye to other un-melted regions on the DNA molecule. Saturating fluorescent dyes, such as EvaGreen and SYTO 9, cannot translocate as the PCR product is already saturated with dye, resulting in a drop in the fluorescent signal. This is shown by the accumulation of dye (black arrows and dark green molecules) outside the DNA molecule. (Reproduced from Pereira et al. 2018).

Compared to SYBR Green dye, both the SYTO9 and EvaGreen dyes have ranked highly in dye performance studies (Varga and James 2006; Li et al. 2010; Eischeid 2011; Radvanszky et al. 2015; Pereira et al. 2018). The SYTO family of fluorescent dyes were originally developed as nucleic acid stains for labelling live cells (Monis et al. 2005; ThermoFisher 2014). In 2005 Monis et al. published work where they used the dye for qPCR and melt curve analysis and compared it to SYBR Green. Since then dyes from the SYTO family, including SYTO 9, have been used in qPCR, HRM analysis, and dye performance comparisons (Gudnason et al. 2007; Li et al. 2010; Eischeid 2011; Bester et al. 2012). Recently, the blending of the SYTO 9 fluorescent stain with other commercial fluorescent dyes demonstrated enhanced qPCR and melt curve performance (Jansson et al. 2017). However, in this study, the BioRad Sso Fast EvaGreen supermix demonstrated higher sensitivity than the QuantaBio AccuMelt supermix with SYTO 9 dye.

The superior sensitivity of the BioRad Sso Fast EvaGreen dye was evidenced by the larger temperature difference between the two LNYV subgroups in the HRM melt peak analysis (Section 2.3.3.5) compared to the overlapping HRM profiles shown with AccuMelt (Section 2.3.3.2). Other studies have also found EvaGreen to be more sensitive, and it has been ranked among the best performing dyes in several studies (Mao et al. 2007; Li et al. 2010; Eischeid 2011; Radvanszky et al. 2015). EvaGreen has also been used successfully in HRM analysis of SNPs in plants (Li et al. 2010), detecting variants of citrus HSVd (Loconsole et al. 2013), detecting molecular markers for PVY resistance (Nie et al. 2016), and the detection of various grapevine (Aloisio et al. 2018) and potato viruses (Cheng et al. 2013).

The Ajithkumar Primers were tested for their suitability in diagnosing the two LNYV subgroups using RT-qPCR-HRM with the BioRad Sso Fast EvaGreen dye. Although the EvaGreen dye proved more sensitive than the AccuMelt dye, these primers were found to be unsuitable for subgroup diagnosis using this approach. These primers did not provide sufficient discrimination between the two LNYV subgroups in a multiplex RT-qPCR-HRM assay, and there was overlap of the HRM melt peaks (Section 2.3.3.5). A temperature difference of 2 °C is considered appropriate for HRM discrimination, although even a difference of 1 °C would still have been acceptable (Cheng et al. 2013; Osathanunkul et al. 2016). The Ajithkumar Primers generated only 0.13 °C between the HRM melt peaks for each subgroup.

Three primer sets that were specific for two-step RT-qPCR-HRM analysis were designed. Although Primer Set 2 appeared suitable with a predicted melt temperature difference of 2°C between the two subgroups (Section 2.3.4), testing revealed this primer set produced the greatest variation among the technical replicates, and the average HRM melt peak temperatures for the two LNYV subgroups differed by only 0.2 °C. This is well below the desired temperature difference of 2 °C (Cheng et al. 2013; Osathanunkul et al. 2016). This may have been due to the larger PCR product size amplified by this primer set. Although the PCR products for each subgroup were < 300 bp (143 bp for subgroup I, and 258 bp for subgroup II), they were the largest PCR products of the three primer sets. Recommended PCR product size does vary within the < 300 bp range, depending on the HRM application. Recent successful virus diagnosis via HRM methods used smaller PCR products. For example, diagnosis of CNRMV and CGRMV generated a PCR product size of 107 bp and 80 bp respectively (Komorowska et al. 2014). Detection of RRV, and variants of citrus HSVd used PCR product sizes of 104 bp, and 62 – 64 bp, respectively (Loconsole et al. 2013; Dobhal et al. 2016). Thus, the PCR products amplified by Primer Set 2 are on the larger side, which may have influenced the usefulness of these primers for LNYV subgroup diagnosis by this approach.

Primer Sets 1 and 3 produced more promising results compared to Primer Set 2. Both Primer Set 1 and 3 amplified subgroup specific products that were distinguishable by HRM melting temperature differences of > 2 °C. While also displaying little variation across the technical replicates. Both were used for diagnosing the LNYV subgroup of samples pre-diagnosed with LNYV, but had not been subgroup typed (Section 2.3.4.5). These primer sets established clear clusters of samples identifying their LNYV subgroup and produced narrow temperature ranges. However, Primer Set 3 demonstrated higher flexibility and sensitivity, appearing to detect subtle differences in the sample sequences, while still diagnosing between subgroup I and subgroup II.

The flexibility and sensitivity of Primer Set 3 was evidenced by the inconsistent grouping of the technical replicates. Group separation appeared to be due to the HRM profile shape, rather than differences in the HRM melt peak temperatures of the technical replicates. The HRM profile shapes were influenced by nucleotide differences in the sequence that had only a limited effect on the melting shape and melt peak temperatures. This would most likely be from class III, C to G, or class IV, A to T, complementary nucleotide differences. Although these complementary nucleotide differences cause only a small impact on the melting shape and temperature (Applied Biosystems 2009; Yamagata et al. 2018), that small impact is more detectable when the PCR product size is also small (Taylor et al. 2011; Słomka et al. 2017), like it is for Primer Set 3 (73 bp for both subgroups).

The flexibility of Primer Set 3 to detect these nucleotide differences could be due to the degenerate bases used for the subgroup I forward primer, or the position of this primer set on the LNYV N gene. The samples used in this analysis had not been sequenced at the time of primer design. Examining these sequences, or the sequences of the PCR products, would indicate if complementary nucleotide differences were impacting the HRM melting profiles (Słomka et al. 2017), and is discussed further later in this discussion.

Furthermore, the subgroup diagnosis assay using Primer Set 3 identified three samples as co-infected with both subgroups, compared to only one of these samples identified by Primer Set 1. Further testing of all three samples using the Ajithkumar Primers, followed by gel electrophoresis, confirmed co-infection of both LNYV subgroups. Even samples that showed late qPCR amplification and humps, rather than peaks, with the Ajithkumar Primers on the HRM melt peak analysis showed PCR products on the gel electrophoresis for both subgroups. This suggests that all samples with irregular HRM melt peak results should be further analysed to confirm the subgroup typing. Further analysis should include confirmation with the Ajithkumar Primers, as performed in this study, but could also include sequencing of the PCR

products (Słomka et al. 2017). In light of this observation, samples SF2 and SF3, which produced a right slope and a left shoulder, respectively, with Primer Set 3 (Section 2.3.4.5.2), should also have been included in confirmation testing or further analysis. The inclusion of artificial positive controls for co-infected infections, or spiked samples, could be used in addition to positive controls of known LNYV subgroup. These would add further support to HRM sample clusters, and aid interpretation of profile shapes (Bester et al. 2012; Jeong et al. 2012; Cheng et al. 2013; Loconsole et al. 2013; Komorowska et al. 2014; Dobhal et al. 2016; Nie et al. 2016; Słomka et al. 2017; Aloisio et al. 2018).

Unexpectedly, late amplification of the subgroup II (HV18) technical replicates was seen with the subgroup I specific primers for Primer Set 3 (Section 2.3.4.3.1) and the Ajithkumar Primers (Section 2.3.4.6). Amplification with the Ajithkumar Primers was also confirmed with a faint band upon gel electrophoresis. It was suspected that perhaps unspecific primer binding was occurring. This could be due to nucleotide mismatches at the 3' end of the primer (Simsek and Adnan 2000; Lorenz 2012). Although unspecific primer binding was not observed with the Ajithkumar Primers using end point RT-PCR (Ajithkumar 2018), the higher efficiency of qPCR could have contributed, making the subgroup I primers more susceptible to mismatching. PCR efficiency can be influenced by the analysis instrument and reagents in the PCR supermix (Bustin et al. 2009; Svec et al. 2015). In particular, the magnesium concentration generated from magnesium chloride ($MgCl_2$) or magnesium sulfate ($MgSO_4$) in the PCR reaction mixture affects both the PCR efficiency and DNA melting behaviour in HRM analysis (ThermoFisher [date unknown]; Applied Biosystems 2009). Ajithkumar (2018) subgroup typed samples using the Bibby Scientific Techne TC-512 Gradient Thermal Cycler and the Invitrogen SuperScript III One-Step RT-PCR with Platinum *Taq* DNA polymerase kit. While in this study, the Roche LightCycler 96 System and the BioRad Sso Fast EvaGreen Supermix was used. Unfortunately, the EvaGreen supermix is proprietary (Mao et al. 2007; Radvanszky et al. 2015) and the component concentrations of the supermix were unavailable to be compared directly with the Invitrogen SuperScript kit used by Ajithkumar (2018). Therefore, the potential contribution of the PCR supermix to the PCR efficiency and unspecific primer binding could not be assessed.

Analysis performed in Chapter 3 provided N gene sequences to compare primer binding sites. This was assessed for the subgroup II (HV18) sample and the three samples showing LNYV subgroup co-infection (RPC1, RPC2 and SF1). The subgroup I forward primer from Primer Set 3 showed the desired six nucleotide differences between the primer and the subgroup II (HV18) sample, which should indicate subgroup specificity (Lorenz 2012; Ajithkumar 2018). Work from Simsek and Adnan (2000) found a G to T mismatch at the 3' end allowed primers to perform as efficiently as the primer to the original sequence. This mismatch could be possible with the 3'

degenerate Y (T/C) base of the subgroup I forward primer of Primer Set 3 with the G nucleotide of the subgroup II (HV18) sample. However, this does not account for the remaining five nucleotide mismatches in the primer, which conveys the primers specificity to subgroup I (Ajithkumar 2018). This is also true of the Ajithkumar subgroup I specific primer, which showed the desired 6 nucleotide differences against the subgroup II (HV18) sequence and was vigorously tested for subgroup specificity by Ajithkumar (2018).

For further clarification of the primer binding sites, sequencing of just the PCR product could be performed. This would provide a clearer picture of the specific sequences that were amplified. For other analyses using HRM analysis to distinguish plant viruses sequencing of the PCR products was performed on randomly chosen samples (Li et al. 2010; Loconsole et al. 2013; Komorowska et al. 2014; Nie et al. 2016). Alternatively, sequencing of the cDNA prior to qPCR-HRM analysis was also performed in the literature. This would provide sequences to further review primer binding sites and perform *in silico* predictive melt analysis using online tools, such as uMelt (Chomič et al. 2011; Downey 2014; Simko 2016; Aloisio et al. 2018). Although additional sequencing was not performed in this study due to the required extra time and costs for samples to be sent offsite for the sequencing, this analysis could provide further confidence in the identification of the three co-infected samples and the sample (HV18) reported to be LNYV subgroup II (Ajithkumar 2018).

Examination of the primer positions against the N gene sequences for the three co-infected samples showed complementarity with the subgroup I primers, but not the subgroup II. This suggests that subgroup II sequences were present in the samples and amplified in the two-step RT-qPCR process, but not in the one-step RT-PCR performed in Chapter 3. This is similar for the subgroup II (HV18) sample. This sample was subgroup typed by Ajithkumar (2018) using one-step RT-PCR, and produced an N gene sequence correlating to other subgroup II samples (Chapter 3), but showed late amplification with the subgroup I specific primers after two-step RT-qPCR. Sequencing of the cDNA would be an appropriate step to confirm the presence of additional sequences in the samples, but was beyond the scope of this study due to time constraints.

The indication of other viral sequences present in the samples suggests a quasi-species nature of LNYV, which is common for RNA viruses. A quasi-species exists as a cloud of closely related, but not identical, viral sequences within the one host (Compans et al. 2006; Viruses are models... 2018). To date, a quasi-species model for LNYV has not been reported. It has, however, been reported for other plant viruses (Section 1.2.3.4; (Schneider and Roossinck 2001; Ali and Roossinck 2017; Jo et al. 2017). Recently co-infection of two variant strains has

been reported for two plant rhabdoviruses, OFV and SCV (Koloniuk et al. 2018; Roy et al. 2020). The co-infection of two SCV strains belonging to SCV subgroup I (Koloniuk et al. 2018) suggests a quasi-species nature for this cytorhabdovirus. In this study, the variation seen in the HRM melt peak temperatures across technical replicates of the same biological sample, also indicate a quasi-species nature for LNYV. As some sample replicates showed no melt peak temperature variation, this would suggest those that did, were due to sequence differences, rather than variation from the instrument's sensitive detection equipment (Roche Diagnostics 2012). In the literature, cloning followed by sequencing has been performed to confirm the quasi-species nature of the virus studied (Benmansour et al. 1997; Schneider and Roossinck 2001; Ali and Roossinck 2017; Díaz-Martínez et al. 2018). Specific sequencing of sample cDNA or the PCR products may have provided data to confirm if the technical replicates of the LNYV samples from this study, were very closely related, but not identical, and therefore suggestive of a quasi-species model for LNYV within its two subgroups. As discussed, analysis of this nature was beyond the scope of this study, but presents an interesting avenue for further research.

Throughout the qPCR experiments the subgroup II sample appeared to show greater PCR efficiency. In multiplex analysis with the Ajithkumar Primers and Primer Set 3, the subgroup II sample (HV19 and HV18) amplified earlier than the subgroup I sample (HV27 and HV14), regardless of the type of fluorescent dye tested in this study (Sections 2.3.2.1, 2.3.3.2 and 2.3.3.5). The subgroup II sample (HV18) also showed greater efficiency with Primer Sets 1 (Section 2.3.4.1.2) and 2 (Section 2.3.4.2.2). This was indicated by the steep exponential phases during amplification (Nolan et al. 2013), and greater fluoresce was seen on the HRM melt peaks, with the subgroup II sample generating taller peaks (Taylor et al. 2011; Nolan et al. 2013). This is suggestive of the subgroup II samples containing a higher titre of viral RNA, as the same total RNA volume was used in all analyses. A higher viral titre could be due to greater virulence of subgroup II compared to subgroup I. This has also been hypothesised as the underlying reason that subgroup II was able to outcompete subgroup I in Australia (Higgins et al. 2016b). Analysis including a validated reference gene for LNYV is needed to accurately measure and normalise the amount of viral titre in the samples and therefore compare subgroup I to subgroup II (Nolan et al. 2013). This was beyond the scope of the current study, as the focus was on diagnosis of the two subgroups. However, these observations provide support for this hypothesis and indicate further investigation into this is appropriate.

Optimisation of the qPCR-HRM protocol could eliminate the unspecific binding of the subgroup II (HV18) sample with the subgroup I primers, and potentially reduce the HRM melt peak temperature variation and profile shape variation seen. Optimisation areas include primer

concentration and sensitivity, and performing gradient RT-qPCR to optimise primer annealing temperature (Roche 2008; Taylor et al. 2011; Lorenz 2012; Nolan et al. 2013). Optimisation of the thermal cycling conditions, number of qPCR cycles, and the magnesium concentration to reduce potential unspecific primer binding (Applied Biosystems 2009; Taylor et al. 2011; Lorenz 2012; Nolan et al. 2013). In addition, determining the analytical sensitivity and efficiency of the assay means the lower limit of detection for samples can be calculated (Lorenz 2012; Nolan et al. 2013). Optimising the HRM protocol by adjusting the instrument temperature ramp rate or the normalisation sensitivity from the default settings could reduce HRM melt peak temperature and profile shape variation (Varga and James 2006; Chomič et al. 2011; Cheng et al. 2013; Nolan et al. 2013). Due to time restraints, these optimisation strategies have not been included in this study.

Virus detection plays an important role in increasing the actual crop yield of crops, and the development of an assay to diagnose between the two LNYV subgroups, as described in this Chapter, has allowed the NZ LNYV subgroup population to be monitored and updated. It was hypothesised that recent high crop losses in NZ could be due to an increased presence of LNYV subgroup II. Diagnosis of 18 samples identified only four as subgroup II, 11 as subgroup I, and three as co-infected with both subgroups. As less than half were identified as subgroup II, this suggests LNYV subgroup II is not exhibiting a stronger presence in NZ. However, subgroup diagnosis alone cannot rule out the hypothesis that a new strain from Australia has arrived in NZ, or provide wider insight into the LNYV subgroup relationships, or relationships between the NZ and Australian LNYV populations. As the NZ and Australian LNYV isolates have shown sequence differences (Higgins et al. 2016b; Ajithkumar 2018), analysis of the entire N gene is required to analyse these relationships. Phylogenetic analysis of the two LNYV subgroups, from NZ and Australia, allows these relationships to be assessed and characterised. This analysis is described in Chapter 3.

The findings of Chapter 2 show that LNYV subgroup diagnosis via qPCR-HRM analysis is feasible, as well as specific, flexible, and cost- and time efficient. Primers and intercalating fluorescent dyes designed specifically for qPCR-HRM analysis showed superior results. Primer Set 3 proved to be highly sensitive and specific, showing superiority over an RT-PCR approach, and the other qPCR-HRM primers developed during this study. Primer Set 3 allowed sequence differences to be identified, along with the LNYV subgroup, as well as identifying three samples as co-infected. This sensitive analysis indicated the presence of multiple sequences within the one biological sample, suggestive of a quasi-species nature for LNYV. Further analysis of these samples is required to provide robust validation of this, and was beyond the scope of this

study, but highlights an avenue for future research. To date this is the first instance of LNYV presenting a quasi-species nature, as well as co-infection of its two subgroups in the one host.

Chapter 3

Phylogenetic Analysis of the LNYV N Gene

3.1 Introduction

In Chapter 2, a two-step RT-qPCR-HRM assay to diagnose between the two LNYV subgroups is described. Three primer sets were developed for LNYV subgroup diagnosis, and Primer Set 3 (LNYVNS1F_844/ LNYVNS1R_916/ LNYVNS2F_844/ LNYVNS2R_916), demonstrated the highest flexibility and sequence sensitivity. This primer set was used to diagnose the LNYV subgroup of 18 previously untyped samples collected in NZ. Of these samples, 11 were diagnosed as subgroup I (WHG1-5, RPO1-4, RPE2, and SF3), four as subgroup II (RPC3, RPE1, RPE3, and SF2), and three as co-infected with both subgroups (SF1, RPC1, and RPC2). These results allowed the NZ LNYV subgroup population structure to be updated.

The low number of subgroup II samples that were identified indicated that an increased presence of subgroup II was unlikely to be the cause of increased crop losses experienced recently in NZ. However, these results cannot rule out the alternative hypothesis that these losses could be due to a new LNYV strain from Australia arriving in NZ (Higgins et al. 2016b; Fletcher et al. 2019). Such a stain could demonstrate a T_m and HRM melt peak that is similar to the NZ subgroup I and subgroup II isolates. Assessment of the taxonomic and evolutionary relationships of these isolates with other published LNYV isolates from NZ and Australia may provide information to evaluate this hypothesis further.

Phylogenetic analysis is an important bioinformatics tool that allows the taxonomic and evolutionary relationships to be assessed and characterised. To examine these relationships for LNYV, phylogenetic analysis of the entire LNYV N gene is required. The LNYV N gene has been used previously to differentiate between the two LNYV subgroups (Callaghan and Dietzgen 2005; Higgins et al. 2016b), and sequence differences have been shown between NZ and Australian LNYV isolates (Higgins et al. 2016b; Ajithkumar 2018). Phylogenetic analysis of the entire N gene of the 18 samples subgroup typed in Chapter 2, with other published LNYV isolates, is described in this Chapter.

3.1.1 Plant Rhabdovirus Phylogenetic Analyses

For taxonomic characterisation of new virus species, phylogenetic analysis of the entire L gene has been used extensively (Pagán 2018; Walker et al. 2018; ICTV 2020). Members of the Rhabdovirus family are categorised into the twenty-two different genera by this type of phylogenetic analysis, among other criteria (Walker et al. 2018; Dietzgen et al. 2020). However, for characterisation of within species relationships, the N gene has been assessed for variability and has been used for phylogenetic analysis for several rhabdoviruses, across the

different rhabdoviral genera (Revill et al. 2005; Pappi et al. 2016; Ramalho et al. 2016; Samarfard et al. 2018; Roy et al. 2020).

Variability and within species relationships have been assessed using the N gene for several alphanucleorhabdoviruses (previously nucleorhabdoviruses). For example, using partial N gene sequences that represented about 75 % of the TaVCV N gene, isolates from the Pacific Islands showed maximum variability values of 19.3 % at the nucleotide level and 6.3 % at the amino acid level (Revill et al. 2005). In addition, the first report of TaVCV in Hawaii showed similar N gene nucleotide variability to the Fijian type strain of TaVCV (Long et al. 2014). Phylogenetic analysis of the partial N gene sequences showed that the isolates generally grouped by geographical location (Revill et al. 2005). A geo-spatial connection between isolates has also been shown for another alphanucelorhabdovirus, EMDV (Pappi et al. 2016). Analysis of full-length N gene sequences showed phylogenetic groupings of EMDV isolates by geographical region. In contrast to TaVCV isolates, the EMDV isolates showed low genetic diversity overall, with the N gene showing up to 99 % sequence identity at the nucleotide and amino acid level (Pappi et al. 2016). These findings are similar to those seen with dichorhaviruses. CoRSV has shown low genetic diversity, but strong geo-spatial relationships after phylogenetic analysis with partial N gene sequences (Ramalho et al. 2016). While the dichorhavirus OFV has shown only 83 % to 86 % nucleotide sequence identity of the N gene, and 97 % at the amino acid level, across its subgroups (Roy et al. 2020). Phylogenetic analysis has also shown the four subgroups of OFV divide based on the plant host. Subgroups I and II (OFV-orc1 and OFV-orc2) infect orchids (Kubo et al. 2009), while subgroups III and IV (OFV-cit1 and OFV-cit2) infect citrus (Roy et al. 2020). Unlike CoRSV, EMDV, and TaVCV, the OFV orchid subgroups (I and II) showed no geographical link (Kubo et al. 2009; Kondo et al. 2017; Roy et al. 2020). Currently the OFV citrus subgroups (III and IV) have only been reported from citrus plants from the same region in Mexico (Roy et al. 2020). Of interest though, is that several citrus plants from both collection sites in Mexico, were found to be co-infected with both OFV citrus subgroups (Roy et al. 2020).

Studies of variability within cytorhabdoviruses have shown both high and low degrees of similarity. ADV isolates from different regions in Argentina showed 99.0 % to 99.9 % nucleotide identity of their N gene sequences (Samarfard et al. 2018). This was also reflected in the phylogenetic analysis, as the 13 isolates grouped extremely close together although they were geographically diverse (Samarfard et al. 2018). Similarly, the only report of SCV N gene variability showed 94 % identity at the amino acid level (Koloniuk et al. 2018). However, phylogenetic analysis of the L gene, showed the two SCV variants both belonged to the same SCV subgroup, and they were both found in the same host plant (Koloniuk et al. 2018). In contrast to ADV and SCV, analysis of two RVCV variants showed high genetic diversity with only

75 % pairwise similarity of the N gene at the amino acid level (Jones et al. 2019). Phylogenetic analysis was carried out on all the viral genes, confirming the studied virus to be a variant of RVCV (Jones et al. 2019).

3.1.2 LNYV Phylogenetic Analysis

For LNYV, variability and phylogenetic analysis of the N gene first identified its two subgroups (Callaghan and Dietzgen 2005). Since then, sequence diversity between, and within, the subgroups has been shown (Higgins et al. 2016b; Ajithkumar 2018). Analysis of NZ LNYV isolates included only two NZ subgroup II samples. Furthermore, all the NZ samples came from one lettuce farm in the North Island and were all collected on the same day (Higgins et al. 2016b; Ajithkumar 2018). Chapter 2 describes the subgroup diagnosis of 18 new samples from both the North and South Islands of NZ; phylogenetic analysis that includes these new samples would provide further insight into LNYV subgroup relationships, and relationships between the NZ and Australian LNYV populations.

The LNYV subgroup phylogenetic analysis carried out previously showed distinct topographic characteristics between the two subgroups. LNYV subgroup I has shown longer branches that are well supported with high bootstrap values, while subgroup II, has shown short branches with low bootstrap values (Higgins et al. 2016b; Ajithkumar 2018). From these data, Higgins et al. (2016) suggested that subgroup II is undergoing rapid dispersal. The phylogenetic analysis performed by Ajithkumar (2018) reflected this hypothesis as well. The rapid dispersal of subgroup II supports the hypothesis that subgroup II has been able to supplant subgroup I in Australia, possibly due to more optimal plant and/or insect host interactions, allowing it to outcompete subgroup I (Higgins et al. 2016b). Although the subgroup diagnosis described in Chapter 2 suggests subgroup II is not outcompeting subgroup I in NZ at this time, a new strain from Australia arriving in NZ could not be ruled out, and this could still supplant subgroup I, and possibly subgroup II, in NZ. Phylogenetic analysis including these new NZ samples, from both the North and South Islands, allows for more insight into the LNYV subgroup and population relationships. This analysis was carried out and is described in this Chapter.

3.1.3 Aims and Objectives

This Chapter describes the phylogenetic analysis of the entire N gene from 43 LNYV isolates from NZ and Australia. This analysis investigates the taxonomic and evolutionary relationships

of these isolates, providing insight into the subgroup population structure of the NZ LNYV isolates and their relationships to each other and the Australian LNYV population. Unpublished NZ LNYV samples, which were subgroup diagnosed as described in Chapter 2, were analysed with other published LNYV N gene sequences. This was achieved by addressing the following specific objectives:

1) Extract and amplify the entire LNYV N gene from unpublished isolates using endpoint RT-PCR.
2) Sequence the unpublished LNYV N gene sequences via Sanger sequencing.
3) Curate the LNYV N gene sequences.
4) Construct ML phylogenetic trees with 1,000 bootstrap replicates using both the nucleotide and translated protein sequences.
5) Analyse the taxonomic and evolutionary relationships inferred by the phylogenetic trees.

3.2 Materials and Methods

3.2.1 Plant Materials

NZ lettuce leaf material known to be infected with LNYV was used for this analysis. This included samples collected in 2018 that were subgroup typed as described in Chapter 2, and two samples collected in 2011 that were subgroup typed by Ajithkumar (2018). The samples collected in 2011 (HV14 and HV18) were provided by Colleen Higgins (Auckland University of Technology, NZ), and the samples collected in 2018 were provided by John Fletcher (formerly of The Institute for Plant & Food Research, NZ). These samples are displayed in Table 3.1, they are grouped by their subgroup and show isolate names, sample location and year of collection.

Table 3.1 NZ LNYV isolates from lettuce leaf material used in this study. Isolate names, locations sampled from, year sampled, and subgroup are shown.

Isolate	Location	Year Sampled
Subgroup I		
HV14	Harrisville, Auckland, North Island	2011
WHG1	Woodhaven Gardens, Levin, North Island	2018
WHG2	Woodhaven Gardens, Levin, North Island	2018
WHG3	Woodhaven Gardens, Levin, North Island	2018
WHG4	Woodhaven Gardens, Levin, North Island	2018
WHG5	Woodhaven Gardens, Levin, North Island	2018
RPO1	Richmond Plains, Nelson, South Island	2018
RPO2	Richmond Plains, Nelson, South Island	2018
RPO3	Richmond Plains, Nelson, South Island	2018
RPO4	Richmond Plains, Nelson, South Island	2018
RPE2	Richmond Plains, Nelson, South Island	2018
SF3	Sefton, Mid Canterbury, South Island	2018
Subgroup II		
HV18	Harrisville, Auckland, North Island	2011
RPC3	Richmond Plains, Nelson, South Island	2018
RPE1	Richmond Plains, Nelson, South Island	2018
RPE3	Richmond Plains, Nelson, South Island	2018
SF2	Sefton, Mid Canterbury, South Island	2018
Co-infected		
SF1	Sefton, Mid Canterbury, South Island	2018
RPC1	Richmond Plains, Nelson, South Island	2018
RPC2	Richmond Plains, Nelson, South Island	2018

3.2.2 RNA Extraction and Quality Testing

Total RNA was extracted from frozen plant leaf material (Section 3.2.1) using a Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, USA) as described in Section 2.2.2. Quality testing of the extracted RNA was carried out via spectrophotometry and gel electrophoresis as described in Section 2.2.2.1.

3.2.3 One-Step SuperScript IV RT-PCR Analysis of the LNYV N Gene

One-step RT-PCR was carried out using the Invitrogen SuperScript IV (SSIV) One-Step RT-PCR System to amplify the entire LNYV N gene. This was performed on a total of 20 NZ lettuce samples (Section 3.2.1), using the primer set BCNG1/ BCNG2, which was designed to amplify the entire LNYV N gene (Callaghan and Dietzgen 2005). The BCNG1 primer sequence is as follows: 5'–TCT GGG TAT TGG TTC GGG AAA AGA GTG–3', binding in the intergenic region between the LNYV *leader* sequence and the N gene. The BCNG2 primer sequence: 5'–AGT ATT CAT AAA CTG ATG TGG TTT CTC–3' binds in the intergenic region between the N and P gene. The Invitrogen SSIV kit was used due to the proofreading aspect of the platinum SuperFi DNA polymerase, and the sensitivity of SSIV to synthesise longer targets (Invitrogen 2018).

Reactions were prepared with 12.5 µL of 2X Platinum SuperFi RT-PCR Master Mix, 0.25 µL SSIV RT Mix, and 1.25 µL each of 10 µM BCNG1 and BCNG2 primers. RNA samples (150 – 300 ng) were added and the reactions brought up to 25 µL with nuclease free water. A sample that had previously amplified was included as a positive control and a sample of nuclease free water in place of RNA was included as an NTC. All samples were prepared on ice.

RT-PCR amplification was carried out using a Bibby Scientific Techne TC-512 Gradient Thermal Cycler (Fisher Scientific, England, UK) under the following conditions: cDNA synthesis was carried out at 50 °C for 30 minutes, followed by amplification at 94 °C for 2 minutes. This was followed by 40 cycles consisting of 94 °C for 30 seconds, 50 °C for 30 seconds, and 68 °C for 2 minutes, concluding with a final extension of 10 minutes at 68 °C, and then held at 15 °C. Samples were stored at -20 °C until required.

3.2.4 Gel Electrophoresis of the One-Step RT-PCR Analysis of the LNYV N Gene

To confirm amplification of the N gene from the one-step SSIV RT-PCR, gel electrophoresis was performed. A 1 % agarose/1 x TBE gel was prepared as per Section 2.2.2.3. Each well contained 2 μL of each sample with 2 μL of loading dye. A 100 bp DNA ladder (Solis BioDyne, Estonia) at 2 μL of 0.1 μg/μL was used as a size marker. The gel was electrophoresed for 50 minutes at 75 Volts, and viewed with an Alpha Imager (Protein Simple, California, USA).

3.2.5 Gel Purification of the One-Step RT-PCR Product of the LNYV N Gene

To eliminate any potential contamination that could interfere with the Sanger sequencing process, the PCR products were purified from the agarose gel. A 1 % agarose/1 x TBE gel (Section 2.2.2.3), was loaded with the remaining PCR product (~ 18 μL) with 2 μL of loading dye. A 100 bp DNA ladder (Solis BioDyne, Estonia) at 2 μL of 0.1 μg/μL was used as a size marker, and 2 μL of the positive control sample was also included to confirm size accuracy. The gel was electrophoresed for 40 minutes at 75 Volts.

The gel purification was performed using the ThermoScientific GeneJET Gel Extraction Kit. Using a UV transilluminator (Alpha Imager; Protein Simple, California, USA), the expected PCR product band was visualised, and excised using a sterile scalpel and placed into a pre-weighed 1.5 mL tube. Binding Buffer was added to each tube in a 1:1 ratio of 100 μL of Binding Buffer for every 100 mg of gel. Each tube was then incubated in a dry heat block at 60 °C for approximately 10 minutes, or until the gel slice had dissolved. Tubes were gently mixed by inversion to assist the gel slice melting process. Once dissolved, the tubes were briefly vortexed. Up to 800 μL at a time was then transferred into a purification column provided by the kit, and centrifuged (Eppendorf 5430R Centrifuge, New South Wales, Australia) at room temperature for 1 minute at 14,000 rpm. The flow-through liquid was then decanted. An additional 100 μL of Binding Buffer was added to the column, and again centrifuged at room temperature for 1 minute at 14,000 rpm, and the flow-through decanted. The Wash Buffer was diluted with ethanol before the first use, as per the GeneJET protocol, and 700 μL was added to the columns. The columns were then centrifuged at room temperature for 1 minute at 14,000 rpm, and the flow-through decanted. The empty column was then centrifuged a further 1 minute at room temperature at 14,000 rpm to remove any remaining Wash Buffer. The column was then transferred to a clean 1.5 mL tube. Added to the centre of the column was 25 μL of Elution Buffer followed by centrifugation a final time at room temperature at

14,000 rpm for 1 minute. The eluted liquid contained the purified PCR product, and samples were stored at - 20 °C until required.

3.2.5.1 Spectrophotometry of the Gel Purified LNYV N Gene PCR Product

The quality and concentration of the purified PCR products were assessed via absorbency measurement using the NanoVue (GE Healthcare, Life Sciences) spectrometer and gel electrophoresis. The GeneJET Elution Buffer was used as a reference, and absorbance at wavelengths 230 nm, 260 nm and 280 nm were measured and recorded for each sample.

3.2.5.2 Gel electrophoresis of the Gel Purified LNYV N Gene PCR Product

To confirm the integrity of the purified PCR products, gel electrophoresis was performed. Each well consisted of 2 µL of each purified PCR sample with 2 µL of loading dye. A 100 bp DNA ladder (Solis BioDyne, Estonia) at 2 µL of 0.1 µg/µL was used as a size marker. These were electrophoresed on a 1 % agarose/1 x TBE gel, made as per Section 2.2.2.3, for 50 minutes at 75 Volts. The gel was then viewed with the Alpha Imager (Protein Simple, California, USA).

3.2.6 Sanger Sequencing and Consensus Sequence Assembly of the LNYV N Gene

PCR products were sent for Sanger sequencing by Macrogen Inc. (Macrogen Inc., Seoul, Rep. of Korea) using the BCNG1/ BCNG2 primer set (Callaghan and Dietzgen 2005), as well as a primer set that amplifies an internal region of the LNYV N gene; LNYV_440F/ LNYV_1185R (Higgins et al. 2016b; Ajithkumar 2018). The internal N gene primer sequences are as follows; LNYV_440F 5'–TGA CAC AGA TTC AGA ACA ACT C–3' and LNYV_1185R 5'–GGG ACA ATC CAT CTC CAC TA–3'. The relative locations of these primer sets on the LNYV N gene are illustrated in Figure 3.1.
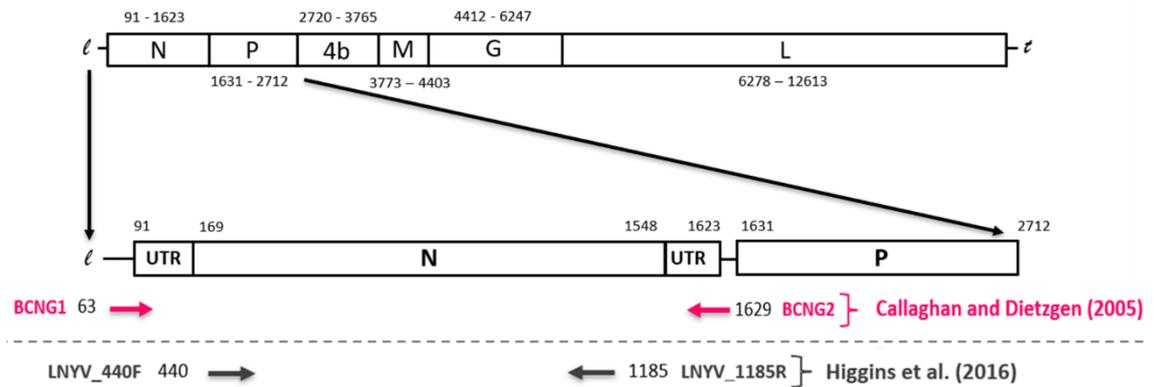
Figure 3.1 Illustration of the relative primer positions for each primer used to sequence the LNYV N gene PCR product. Primer set BCNG1/ BCNG2 (pink; Callaghan and Dietzgen 2005), amplifies the whole N gene by binding in the intergenic regions. The LNYV internal primers (grey; LNYV_440F/ LNYV_1185R; Higgins et al. 2016), amplify a 750 bp internal region of the LNYV N gene. The numbers in black represent the primer binding positions when considering the whole LNYV genome.

Once sequencing was complete, the data was analysed using Geneious v 6.0.6 (www.geneious.com, Biomatters, Auckland, New Zealand). For each sample, four sequence reads were generated from the two primer pairs. Sequences were curated by trimming at both ends to eliminate unreliable or poor-quality base calls. The curated sequences from the two forward primers (BCNGI/ LNYV_440F), and the two reverse primers (BCNG2/ LNYV_1185R), respectively, were pairwise aligned and the alignments visually examined for base call quality and read placement. After further trimming, the forward and reverse contig sequences were multiple aligned to form one consensus sequence. This was also visually examined for placement of sequences and the corresponding base call quality. The consensus sequence for each sample was then extracted for use in downstream bioinformatics analysis.

3.2.7 Phylogenetic Analysis of the LNYV N Gene

Phylogenetic analysis was performed with the N gene of the 20 LNYV isolates sequenced, combined with previously published LNYV N gene isolate sequences from NZ and Australia (Callaghan and Dietzgen 2005; Higgins et al. 2016b). A total of 26 NZ isolates and 17 isolates from Australia were used (Table 3.2). The samples identified as co-infected in Chapter 2 are grouped with the isolates identified as subgroup I. This is due to the relative amounts of subgroup I and subgroup II that were amplified in the RT-qPCR-HRM analysis, where subgroup I appeared more abundant in each case (Chapter 2). The NZ sample locations are illustrated in Figure 3.2, and Australian sample locations in Figure 3.3. The published N gene nucleotide

sequences from two other Cytorhabdoviruses, LYMoV (accession number EF687738) and PeVA (accession number AB735628) were also included as the outgroup for this study. The N gene sequences for LYMoV and PeVA were obtained from the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/).
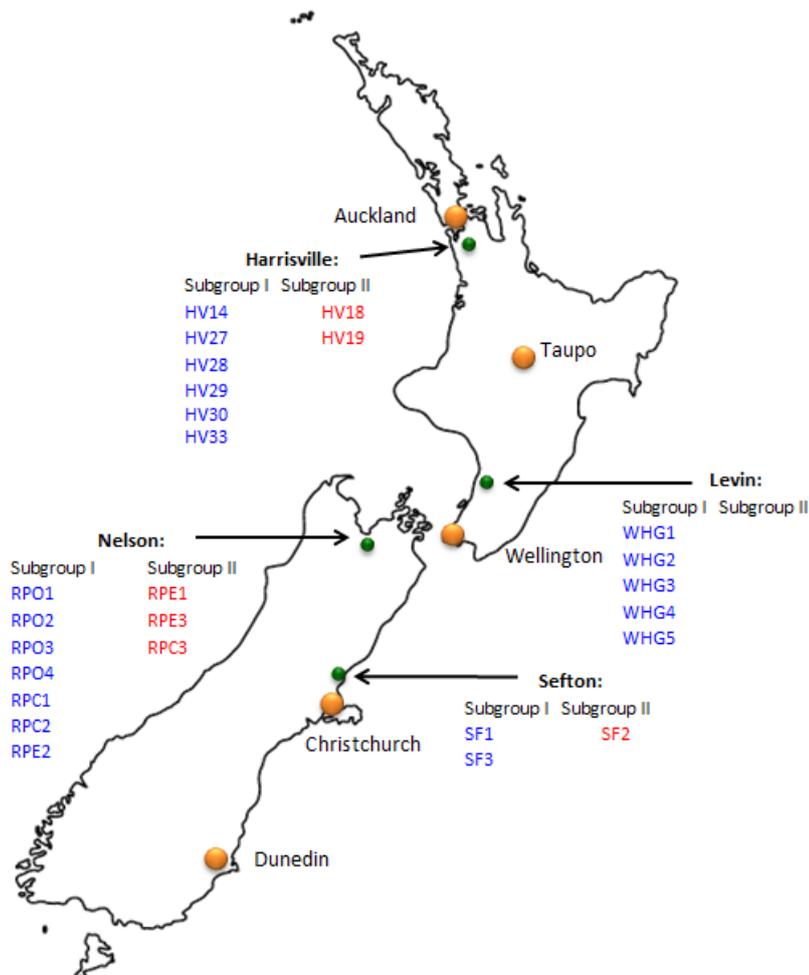
Figure 3.2 Illustration of New Zealand showing the locations of all NZ samples used in phylogenetic analysis. Orange pins indicate major NZ cities, and green pins indicate approximate sample sites. Blue text represents isolates that have been diagnosed as LNYV subgroup I, and red text represents isolates that have been diagnosed as LNYV subgroup II. Isolates diagnosed as co-infected (Chapter 2) have been grouped with subgroup I isolates.

Table 3.2 NZ and Australian LNYV N gene isolates used for phylogenetic analysis. Isolates are grouped by subgroup and country, with isolate name, accession number (where applicable), year and location sampled. Isolates diagnosed as co-infected in Chapter 2 have been included in subgroup I.

Subgroup I				Subgroup II			
Isolate / Published code [a]	Accession number	Year	Location	Isolate/ Published code [a]	Accession number	Year	Location
Australian isolates				**Australian isolates**			
AU2	L30103	1985	Victoria	AU1	AJ746190	1999	Queensland
AU2*	AJ746191	1999	Victoria	AU3	AJ746192	1992	New South Whales
AU6	AJ746195	1993	South Australia	AU4	AJ746793	2000	South Australia
NZ isolates				AU5	AJ746194	2000	Western Australia
HV27/ NZ2 [a]	KP109950	2011	Harrisville, Auckland, North Island	AU7	AJ746196	1993	South Australia
HV28/ NZ3 [a]	KP109951	2011	Harrisville, Auckland, North Island	AU8	AJ746197	1994	South Australia
HV29/ NZ4 [a]	KP109952	2011	Harrisville, Auckland, North Island	AU9	AJ746198	1994	South Australia
HV30/ NZ5 [a]	KP109953	2011	Harrisville, Auckland, North Island	AU10	KP109940	2000	Western Australia
HV33/ NZ6 [a]	KP109954	2011	Harrisville, Auckland, North Island	AU11	KP109941	2001	Gatton, Queensland
HV14	-	2011	Harrisville, Auckland, North Island	AU12	KP109942	2001	Gatton, Queensland
WHG1	-	2018	Woodhaven Gardens, Levin, North Island	AU13	KP109943	1999	Bowen, Queensland
WHG2	-	2018	Woodhaven Gardens, Levin, North Island	AU14	KP109944	2003	Bundaberg, Queensland
WHG3	-	2018	Woodhaven Gardens, Levin, North Island	AU17	KP109947	2002	Mildura, Victoria
WHG4	-	2018	Woodhaven Gardens, Levin, North Island	AU18	KP109948	2005	Goolwa, South Australia
WHG5	-	2018	Woodhaven Gardens, Levin, North Island	**NZ isolates**			
RPO1	-	2018	Richmond Plains, Nelson, South Island	HV19/ NZ1 [a]	KP109949	2011	Harrisville, Auckland, North Island
RPO2	-	2018	Richmond Plains, Nelson, South Island	HV18	-	2011	Harrisville, Auckland, North Island
RPO3	-	2018	Richmond Plains, Nelson, South Island	RPC3	-	2018	Richmond Plains, Nelson, South Island
RPO4	-	2018	Richmond Plains, Nelson, South Island	RPE1	-	2018	Richmond Plains, Nelson, South Island
RPE2	-	2018	Richmond Plains, Nelson, South Island	RPE3	-	2018	Richmond Plains, Nelson, South Island
RPC1	-	2018	Richmond Plains, Nelson, South Island	SF2	-	2018	Sefton, Mid Canterbury, South Island
RPC2	-	2018	Richmond Plains, Nelson, South Island				
SF1	-	2018	Sefton, Mid Canterbury, South Island				
SF3	-	2018	Sefton, Mid Canterbury, South Island				

[a] refers to the simplified published isolate code for NZ isolates (Higgins et al. 2016b). Isolate AU2* refers to isolate AU2 after 10 years of mechanical inoculations in *N. glutinosa* (Callaghan and Dietzgen 2005).

Figure 3.3 Illustration of Australia showing the locations of published LNYV isolates. Isolate AU2* refers to isolate AU2 after 10 years of mechanical inoculations in *N. glutinosa* (Callaghan and Dietzgen 2005). States within Australia have been indicated, orange pins indicate major Australian cities, and green pins indicate approximate sample sites. Blue text represents isolates that have been diagnosed as LNYV subgroup I, and red text represents isolates that have been diagnosed as LNYV subgroup II.

A total of 43 LNYV isolates, and the N gene nucleotide sequences from LYMoV and PeVA were aligned via the MUSCLE algorithm (Edgar 2004) in Geneious v 6.0.6 (www.geneious.com, Biomatters, Auckland, New Zealand). This alignment was visually examined, and curated by trimming ends or removing gaps where required, assisting in the alignment of sequences. The sequences were also translated in the appropriate reading frame, aligned, and trimmed to the N gene ORF in Geneious v 6.0.6. Sequences were imported into MEGA7 (www.megasoftware.net; Kumar et al. 2016) for phylogenetic tree construction. The sequences were realigned by the MUSCLE algorithm in MEGA7. Evolutionary model testing identified the Tamura 3-parameter (T92; Tamura 1992), using a discrete Gamma distribution (+G) and evolutionarily invariable (+I) sites (Kumar et al. 2016) as the most appropriate for the nucleotide sequences. The Jones-Taylor-Thornton (JTT; Jones et al. 1992) +G was identified as most appropriate for the translated protein sequences. The evolutionary models for both sequence types were determined by the lowest Bayesian Information Criterion (BIC) value (Luo et al. 2010). ML trees were constructed for both the nucleotide and protein sequences, each with 1,000 bootstrap replicates for statistical support of the final trees. MEGA7 was used to calculate the pairwise distances for the nucleotide and protein sequences.

3.3   Results

3.3.1 One-Step SuperScript IV RT-PCR Analysis and Purification of the LNYV N Gene

In order to sequence the LNYV N gene, the entire N gene was amplified via one-step RT-PCR using the BCNG1/ BCNG2 (Callaghan and Dietzgen 2005) primer pair. This analysis was performed with 20 lettuce samples from NZ, previously diagnosed with LNYV. Two samples, HV14 and HV18 had been subgroup typed previously (Ajithkumar 2018), but not sequenced. The remaining 18 samples had been subgroup typed as part of this research (described in Chapter 2). Figure 3.4A shows an example agarose gel with sample RPO1, the positive control sample HV33, and an NTC reaction. Amplification of the expected 1,500 bp RT-PCR product was confirmed, and no amplification of the NTC was seen. However, in addition to the expected sized product, primer dimer was observed below the 100 bp marker. This meant the PCR products had to be gel extracted prior to sequencing. Figure 3.4B shows an example gel of the purified N gene PCR products for several samples, confirming the removal of primer dimer. The purified PCR products for all 20 samples were then sequenced.
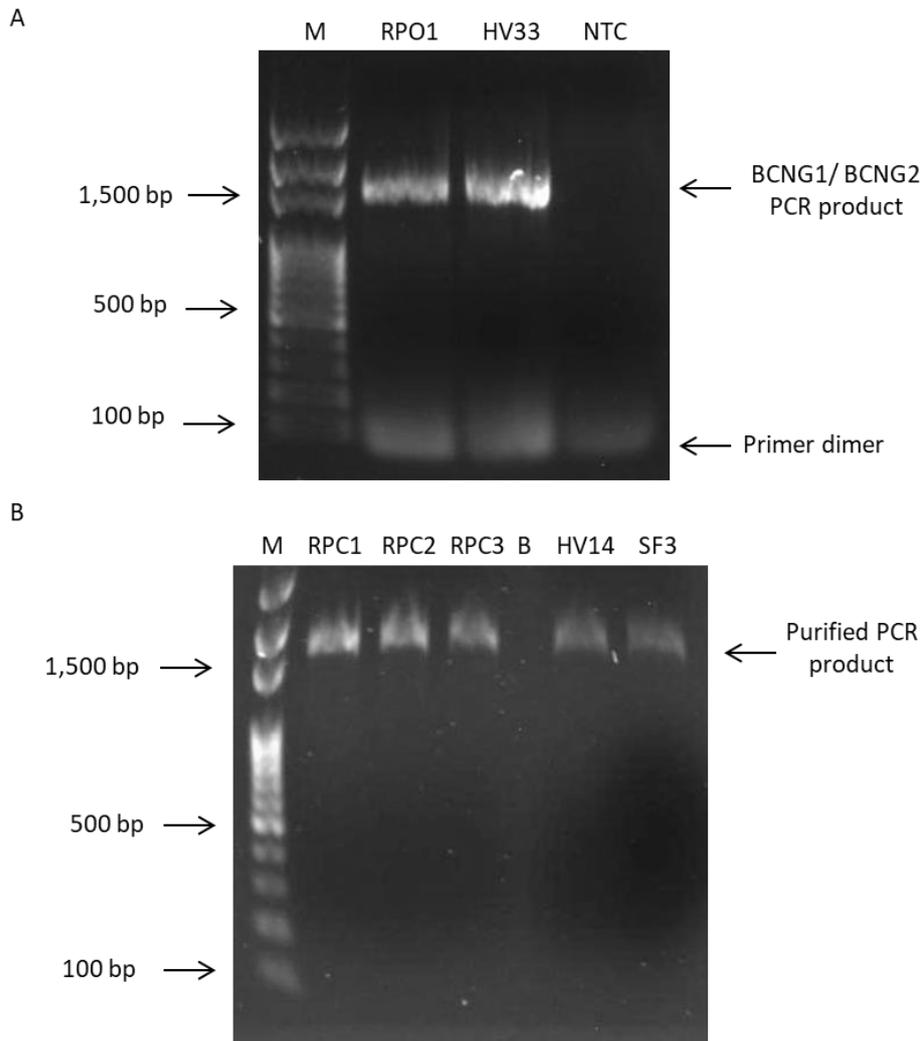
Figure 3.4 Amplifying the LNYV N gene. A) Example gel from one-step SSIV RT-PCR with BCNG1/ BCNG2 primers. Gel shows the expected sized product at 1,500 bp and primer dimer. Lane M = 100 bp DNA ladder, RPO1 = example sample, HV33 = positive control sample, and NTC = no template control. B) Example gel of purified PCR products. Gel shows expected sized product and the removal of primer dimer. Lane M = 100 bp DNA ladder, RPC1, RPC2, RPC3, HV14, and SF3 = example samples. Lane B = Blank, which was deliberately left blank.

3.3.2 Phylogenetic Analysis of the LNYV N Gene

Phylogenetic analysis was carried out using curated sequences for each PCR product and published LNYV N gene sequences to assess their taxonomic and evolutionary relationships. This also allowed the LNYV subgroup population structure in NZ to be evaluated. Figure 3.5 and Figure 3.6 show the ML analysis of the N gene nucleotide sequences of 43 LNYV isolates, LYMoV and PeVA. Figure 3.6 shows a zoomed in, subtree, view of the LNYV branching pattern of Figure 3.5. Figure 3.7 shows the ML analysis of the translated N gene sequences. Both Figure 3.5 and Figure 3.7 show the average pairwise distance values for the LNYV sequences.

The nucleotide phylogenetic analysis (Figure 3.5) showed well supported (82 %) branching of all the LNYV isolates as belonging to LNYV. All LNYV isolates were also separate from LYMoV, the closest outgroup cytorhabdovirus. This is consistent with other studies of the N gene for cytorhabdoviruses (Bejerman et al. 2015; Dietzgen et al. 2017). The LNYV isolates and LYMoV also share a common ancestor.

The LNYV isolates formed two, well supported monophyletic clades. These clades were made up of subgroup I (97 % bootstrap support) and subgroup II (100 % bootstrap support), consistent with previous reports (Callaghan and Dietzgen 2005; Higgins et al. 2016b; Ajithkumar 2018). The LNYV subgroup I clade appeared closer to the last common ancestor, as reported by previous LNYV N gene phylogenetic assessments (Higgins et al. 2016b; Ajithkumar 2018). The majority of NZ isolates belonged to subgroup I, while the majority of Australian isolates belonged to subgroup II. Within each subgroup clade, the NZ and Australian isolates formed well supported, country specific subclades. This has also been reported previously (Higgins et al. 2016b; Ajithkumar 2018), and are indicated on Figure 3.5.

The average pairwise distance values calculated from the nucleotide sequences varied (Figure 3.5). Average pairwise distance between the sequences within the subgroup I clade was low at 0.0134. This indicates high sequence similarity across these isolates, and a low level of sequence diversity. The average pairwise distance within the subgroup II clade was larger at 0.0326. This indicates less sequence similarity, and more divergence, across the subgroup II sequences. The average pairwise distance between subgroup I and subgroup II was 0.278, higher than the subgroups individually, indicating the sequences of the two LNYV subgroups were distinctly different. This supports the phylogenetic topology of two monophyletic clades – one for each subgroup.
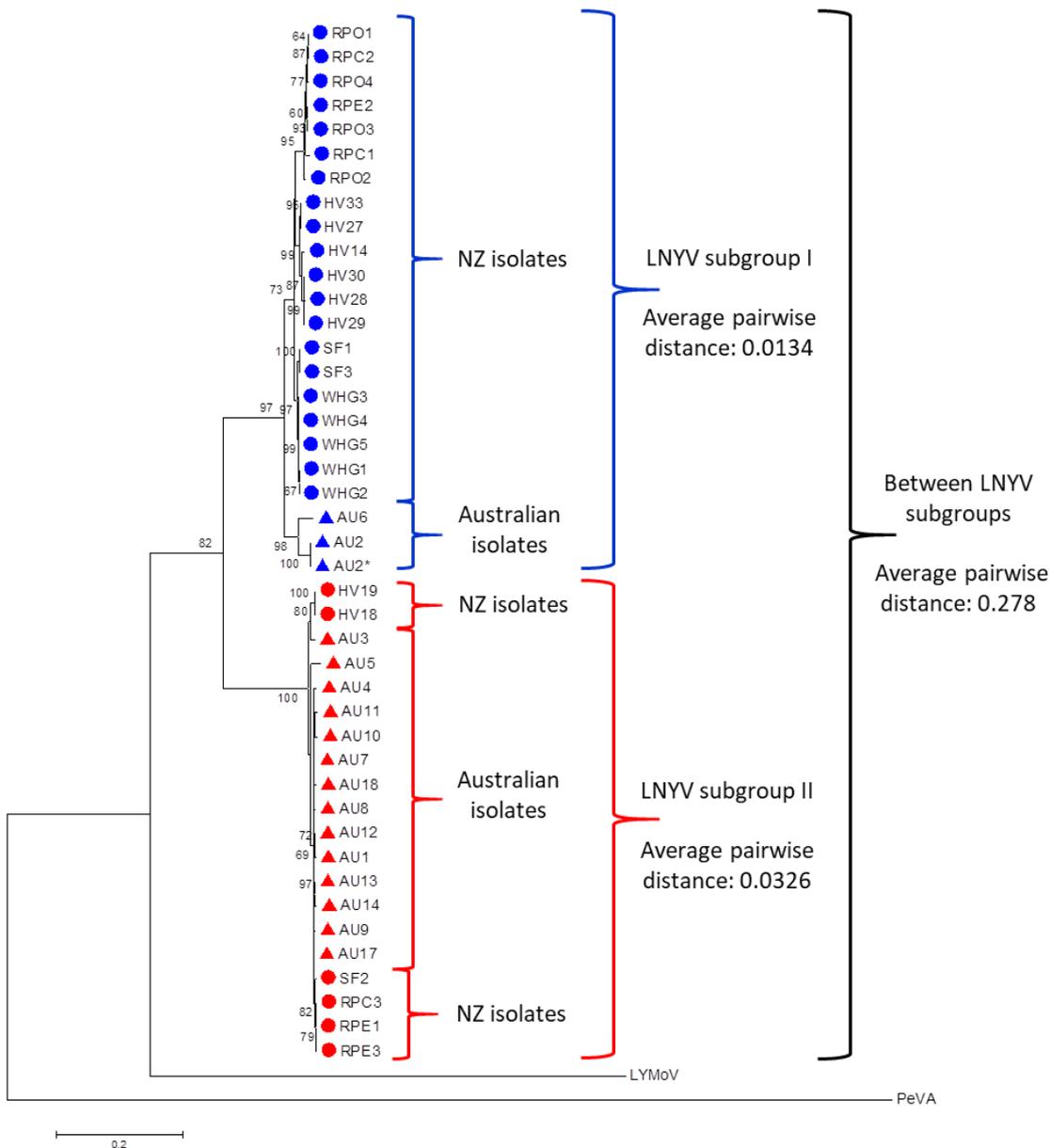
Figure 3.5 Phylogenetic analysis of LNYV N gene nucleotide sequences from isolates from NZ and Australia. LNYV isolates coloured blue for subgroup I, and red for subgroup II. NZ isolates represented with circles, and Australian isolates represented with triangles. Isolate AU2* refers to isolate AU2 after 10 years of mechanical inoculation in *N. glutinosa* (Callaghan and Dietzgen 2005). Outgroup cytorhabdoviruses include lettuce yellow mottle virus (LYMoV) and persimmon virus A (PeVA). ML phylogenetic tree made with 1,000 bootstraps. Bootstrap values > 50 % are shown. The scale represents the number of substitutions per site (0.2).

The topology of the LNYV isolates only (Figure 3.6) was examined using a zoomed in subtree view of Figure 3.5. This was to aid assessment of the taxonomic relationships within each LNYV subgroup clade. Figure 3.6 shows subgroup I bifurcated into two subclades by geographic origin – one of NZ isolates (73 % bootstrap value), and one of Australian isolates (98 % bootstrap value). While the subgroup II clade showed two separate groupings of the NZ isolates.

180

Within the Australian subgroup I subclade (Figure 3.6) the AU2 isolates appeared closer to the subgroup I clade last common ancestor than the AU6 isolate. Only three Australian subgroup I isolates are available due to subgroup I apparently becoming extinct (Callaghan and Dietzgen 2005; Higgins et al. 2016b). Obtaining more subgroup I samples is therefore difficult, and the contribution of sampling error to the number of subgroup I Australian isolates cannot be calculated.

Within the NZ subgroup I subclade (Figure 3.6), isolates appeared to be separated by location. Samples RPO1-4, RPC1, RPC2 and RPE2 clustered together, forming their own nested clade. These samples are all from the Nelson region of the South Island (Figure 3.2). The samples from Harrisville (HV14, HV27, HV28, HV29, HV30 and HV33), in the North Island (Figure 3.2), formed a second nested clade. The third nested clade included samples WHG1-5, from Levin in the North Island, and SF1 and SF3, from Sefton in the South Island (Figure 3.2). All three NZ nested clades were well supported. While the lineage leading to the Harrisville and Nelson isolates appeared just before the lineage of the Levin and Sefton samples, the latter samples showed less branches and therefore less variation between isolates. These isolates also appeared the closest to the last common ancestor of all the subgroup I samples.

The subgroup II branch topology was visually different compared to the subgroup I clade topology (Figure 3.6). Although there was an initial bifurcation into two subclades, similar to the subgroup I topology, from here subgroup II consisted of short branches and low bootstrap values. In addition, the NZ and Australian isolates grouped within both subclades. The majority of the Australian isolates grouped in one subclade, with a polytomy branching pattern with low bootstrap support values indicating rapid dispersal. This also suggests the isolates are rapidly diverging in Australia and supports the findings of Higgins et al. (2016b). This was difficult to infer for the NZ isolates, as only a small number belonged to subgroup II. The NZ isolates appeared to show less divergence from each other, with higher bootstrap values. In general, the subgroup II topology indicates sequence divergence, particularly across the Australian isolates, which was also supported by the larger average pairwise difference value that was calculated for this clade (Figure 3.5).

There were clear country specific isolate groupings across the two subgroup II clades. However, the Australian isolates did not show relationships associated with isolate geographical origin, unlike the NZ isolates in the subgroup I clade. Isolates from across several Australian states grouped together in one subclade, which showed mostly polytomy branches with only a few bifurcating branches and no nested clades. In contrast, the NZ subgroup II isolates did suggest broad geographical groupings. Isolates from the NZ North Island (HV18,

and HV19) grouped in one subclade while isolates from the NZ South Island (SF2, RPC3, RPE1, and RPE3) grouped in the other subclade. As there are limited numbers of NZ subgroup II isolates this could be an artefact of the sampling size.

The branching pattern seen with the subgroup II isolates is suggestive of an Australian origin for this clade. Both subclades include Australian isolates with the NZ isolates. In particular, the NZ South Island isolates share a direct lineage from other Australian isolates. In contrast, the origin of the subgroup I clade is more difficult to determine. The Australian and NZ isolates have grouped separately. Forming two subclades from a bifurcating branch. The limited number of Australian subgroup I isolates available could have an impact on this.
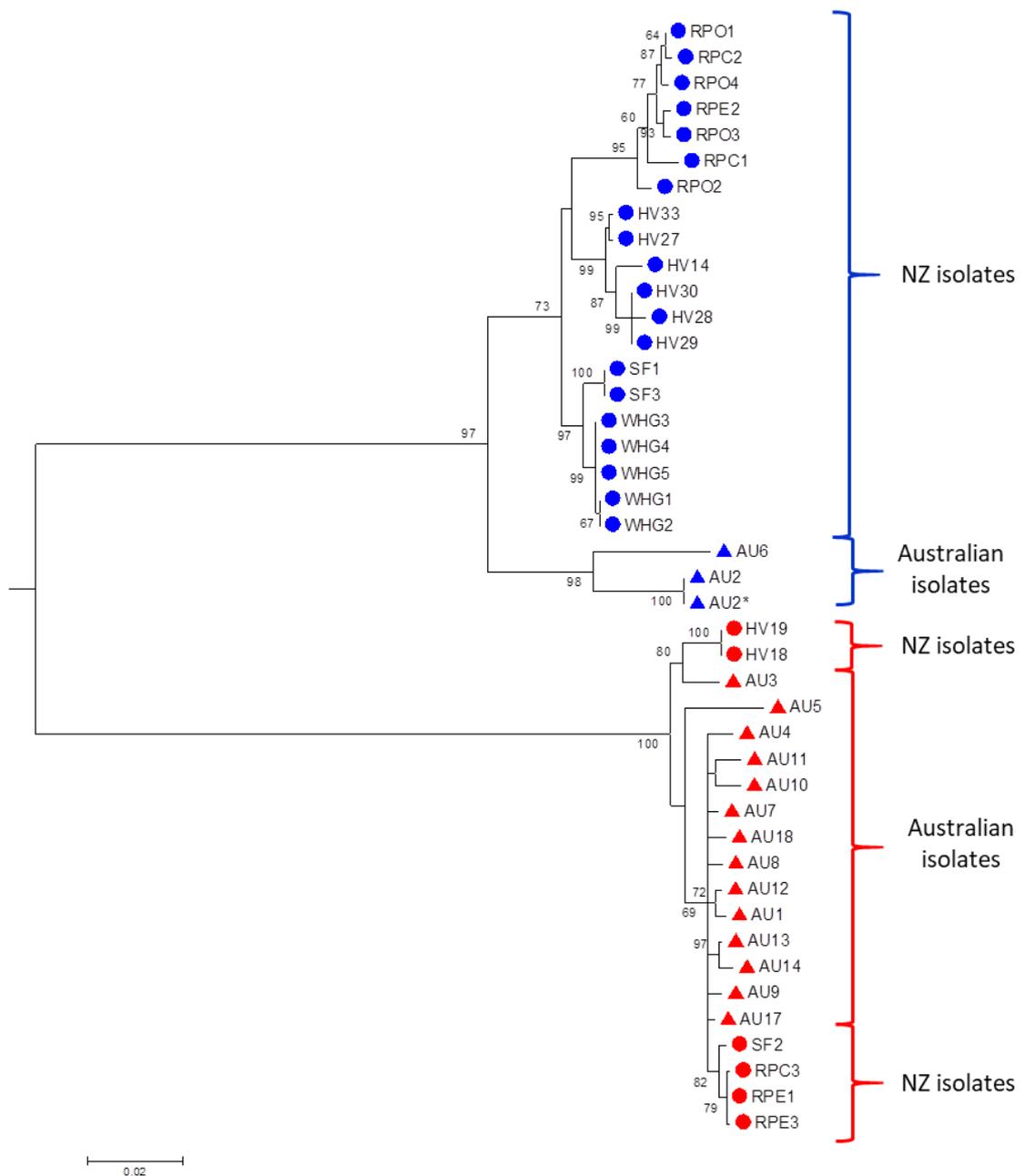
Figure 3.6 Phylogenetic subtree of LNYV N gene nucleotide sequences only from Figure 3.5. LNYV isolates coloured blue for subgroup I, and red for subgroup II. NZ isolates represented with circles, and Australian isolates represented with triangles. Bootstrap values > 50 % are shown. The scale represents the number of substitutions per site (0.02) and differs from the scale shown in Figure 3.5

The phylogenetic tree using the translated protein sequences (Figure 3.7) showed similar topology to the nucleotide tree (Figure 3.5 and Figure 3.6). Two monophyletic clades for LNYV, and country specific groupings within each subgroup clade were observed. As expected, the translated protein tree indicated less divergence overall between isolates compared to the nucleotide tree (Figure 3.5 and Figure 3.6), suggesting greater conservation of the amino acid sequence.

Subgroup I protein sequences showed less divergence compared to that seen in the nucleotide tree. Although the regional groupings of the NZ isolates were maintained, fewer branches were seen within these nested clades. For example, fewer branches were seen with the Nelson isolates (RPC1, RPC2, RPO1 -4, and RPE2), and the Levin isolates (WHG1-5) were identical, which was not the case for the nucleotide tree (Figure 3.5 and Figure 3.6). Branch topology of the subgroup I clade was well supported, suggestive of a more stable population for this subgroup, when compared to subgroup II.

Isolate branching in the subgroup II clade was similar to the nucleotide topology. The NZ isolates were still grouped in the two subgroup II subclades, and the Australian subgroup II isolates showed only slight differences in their positioning compared to the nucleotide tree (Figure 3.5 and Figure 3.6). As with the nucleotide tree, Australian isolates from several states grouped together, showing no link between geographical origin and groupings. Low bootstrap support values were still seen, suggesting rapid dispersal compared to subgroup I. Of interest, the position of the Australian isolate, AU5, suggests the origin of subgroup II is Australian, whereas this relationship is less clear in the nucleotide tree (Figure 3.5 and Figure 3.6).

The average pairwise distances from the protein sequences (Figure 3.7) were lower than those calculated from the nucleotide sequences (Figure 3.5). This indicates conservation of the amino acids that make up the protein sequences. This is also seen in the decrease in the number of branches. The distances between the two LNYV subgroups was 0.029, larger than the subgroup I clade distances of 0.007, and 0.003 for the subgroup II clade (Figure 3.7). The larger pairwise distance values between the two subgroups indicates the protein sequences are distinct from each other, which is also supported by the subgroups forming two well supported monophyletic clades. In contrast to the nucleotide sequences (Figure 3.5), the subgroup II clade showed less pairwise distances than the subgroup I clade.
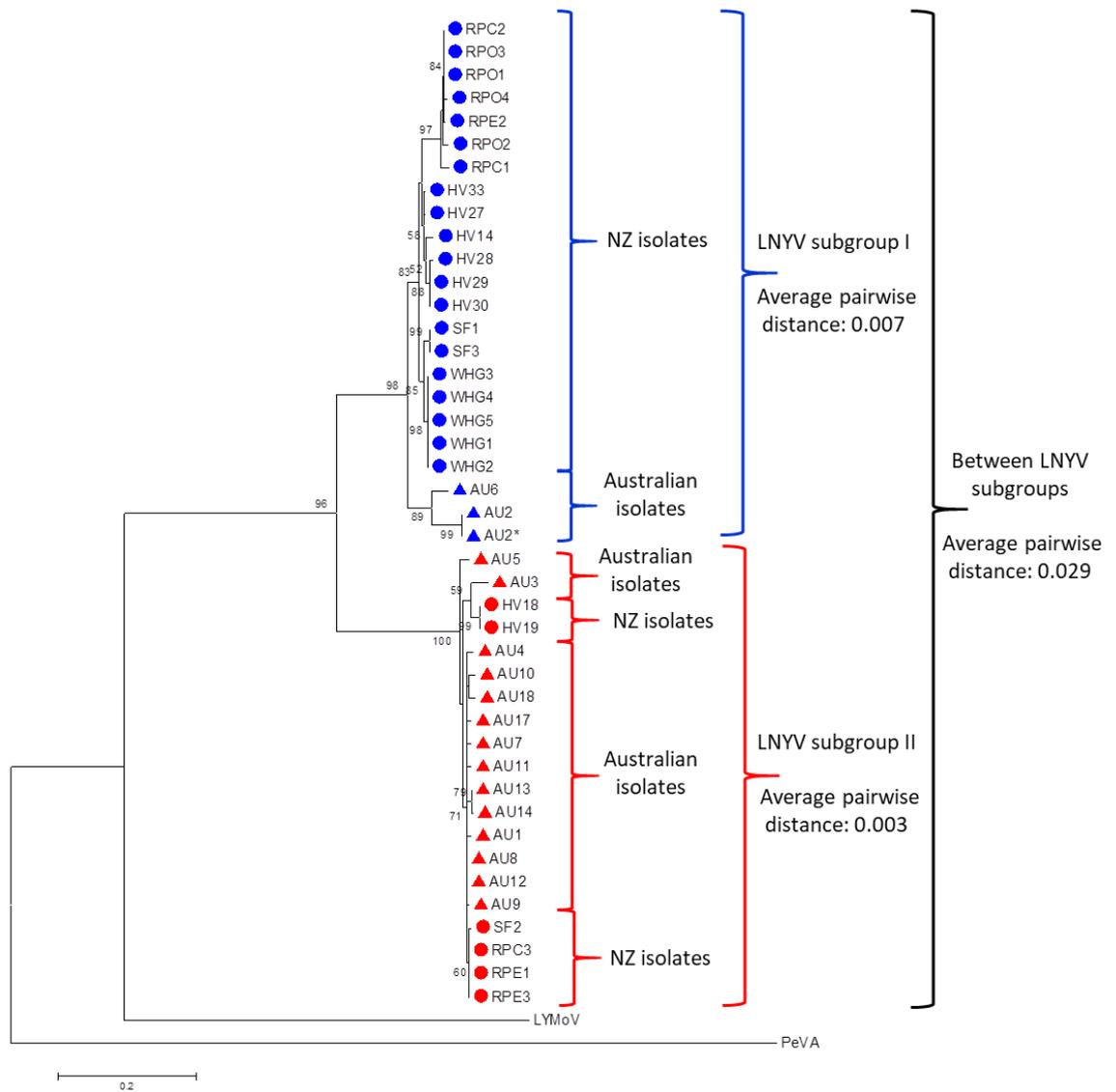
Figure 3.7 Phylogenetic analysis of the translated N gene protein sequences from LNYV isolates from NZ and Australia. LNYV isolates coloured blue for subgroup I, and red for subgroup II. NZ isolates represented with circles, and Australian isolates represented with triangles. Isolate AU2* refers to isolate AU2 after 10 years of mechanical inoculations in *N. glutinosa* (Callaghan and Dietzgen 2005). Outgroup cytorhabdoviruses include Lettuce yellow mottle virus (LYMoV) and Persimmon virus A (PeVA). ML phylogenetic tree made with 1,000 bootstraps. Bootstrap values > 50 % are shown. The scale represents the number of substitutions per site (0.2).

3.4 Discussion

This Chapter describes the phylogenetic analysis of 43 N gene sequences from NZ and Australian LNYV isolates. Analysis of both the nucleotide and protein sequences indicated that LNYV and LYMoV, the closest outgroup cytorhabdovirus, share a common ancestor. The two LNYV subgroups formed two well supported monophyletic clades, where subgroup I emerged earlier. In addition, within both subgroups the NZ and Australian isolates grouped separately by country of origin. These characteristic relationships are consistent with previous phylogenetic analyses (Higgins et al. 2016b; Ajithkumar 2018).

BEAST (https://beast.community) analysis conducted by Higgins et al. (2016b) reported that LNYV subgroup I emerged earlier (150 years ago) than subgroup II (75 years ago). This is also supported by the present study, as both the nucleotide (Figure 3.5) and protein sequence (Figure 3.7) trees showed subgroup I emerged earlier than subgroup II. Ajithkumar (2018) and Higgins et al. (2016) both reported this for their nucleotide analyses. However, their protein trees showed less distinction between the subgroups (Section 1.2.3.3; Figure 1.11). It is possible the addition of more NZ sequences in this study allowed for greater definition of the evolutionary lineage within clades, at the amino acid level. New BEAST or molecular clock analysis with these additional LNYV isolates should be performed to get a better estimate of subgroup emergence dates and their evolutionary relationships. This type of analysis requires substitution and mutation rate data (Lam et al. 2010), of which can be influenced by the host environment (Schneider and Roossinck 2001; Jo et al. 2017). Therefore, this type of analysis was beyond the scope of the current study and is discussed in more detail later in this section.

The data suggest an Australian origin for subgrouop II, but the country specific origin of subgroup I is unclear. Higgins et al. (2016) presents several hypotheses for the origins of LNYV, including that the virus is endemic to NZ and Australia, as first suggested by Randles and Carver (1971). The subgroup I topology clearly bifurcates into the separate NZ and Australian subclades. Additional Australian subgroup I isolates may have provided further resolution for this clade. Unfortunately, acquiring additional isolates is difficult with this subgroup appearing extinct in Australia (Callaghan and Dietzgen 2005; Higgins et al. 2016b). Interestingly, both the nucleotide (Figure 3.5) and the translated protein (Figure 3.7) trees suggest an Australian origin for subgroup II. Both NZ subclades group with Australian isolates, with one subclade appearing to share a direct lineage with Australian isolates. More epidemiological and genomic data is needed to clarify these observations and aid determination of the country specific origins of LNYV.

The topology of the subgroup II clade suggests rapid independent dispersal of the isolates. The Australian isolates did not show groupings by geographic origin, in contrast to the limited NZ isolates. This aspect will be discussed further later in this Section. Short branches were seen for all isolates, with the Australian isolates showing polytomy branching with low bootstrap support values. The pairwise distance value - of the subgroup II clade as a whole - supported these observations. The nucleotide pairwise distance for subgroup II (Figure 3.5; 0.0326) was larger than subgroup I (0.0134), indicating higher sequence variation across the isolates in this subgroup. The protein pairwise distance value for subgroup II (Figure 3.7; 0.003) was smaller than that for subgroup I (0.007), and indicates conservation of the translated proteins. This would suggest nucleotide variations occurred in the third codon position leading to synonymous substitutions, rather than nonsynonymous substitutions (Jordan et al. 2000; Lam et al. 2010; Pagán 2018). This could indicate that negative selective pressure, to remove deleterious or unfavourable variation, was impacting on subgroup II as the encoded proteins and their function were not altered. Negative selective pressure is also referred to as purifying selection to increase the fitness of the viral population (Lam et al. 2010; Pagán 2018).

Negative selective pressure has been reported to strongly influence single-stranded RNA viruses. In particular on coat protein genes, due to their interaction with vectors and hosts (Benmansour et al. 1997; Jordan et al. 2000; Lam et al. 2010; Pagán 2018). This suggests the G gene and its protein may be under negative selective pressure due to its involvement with the aphid host (Pappi et al. 2016; Gallet et al. 2018). Negative selective pressure was reported for all EMDV genes, including the G gene (Pappi et al. 2016), while negative selective pressure was reported for the N genes of ADV (Samarfard et al. 2018) and CoRSV isolates (Ramalho et al. 2016).

The action of negative selective pressure acting on subgroup II could add support to the hypothesis that subgroup II was able to out compete subgroup I in Australia. It has been suggested that subgroup II may have developed more optimal interactions with vectors and/or host plants (Higgins et al. 2016b). Therefore, increasing the viral fitness of the subgroup II population and leading to the extinction of the subgroup I population. Analysis of the G gene and its protein, due to its role in attachment to the aphid vector (Sections 1.2.1.1 and 1.2.3.1), may also provide additional information to assess this hypothesis further (Ajithkumar 2018).

On the other hand, the evolution of subgroup I appeared to be more stable. The phylogenetic topology of subgroup I reflected a more gradual, established evolution, compared to subgroup II. Longer, well supported branches, and nested clades of the NZ isolates were observed. The earlier emergence of subgroup I, which has meant more time has elapsed, has allowed for

sequence variation to become established. The nucleotide pairwise distance value was low (Figure 3.5; 0.0134) and smaller than that for subgroup II, indicating fewer nucleotide differences across the isolates in this subgroup. In contrast, the protein pairwise distance value (Figure 3.7; 0.007) was larger for subgroup I than subgroup II. This could indicate that the nucleotide distances calculated were from nonsynonymous substitutions, and therefore, translating into different amino acids. This is suggestive of the influence of positive selective pressure to increase the viral fitness of the population (Lam et al. 2010; Pagán 2018).

Further *in silico* analysis would provide additional data on the selective pressures impacting both LNYV subgroups. Selective pressure can be measured by counting substitutions, ML methods, or Tajima's test statistic (Lam et al. 2010; Pagán 2018). This type of analysis would provide further insight into the pressures influencing evolution of the isolates within each subgroup. Further differentiating and analysing the isolates by country within each subgroup may help characterise patterns of country specific selective pressures acting on the isolates. Lack of time precluded these analyses being done here, but should be considered in the future.

As well as forming country specific clades within both subgroups, the NZ isolates showed further regional groupings, whereas the Australian isolates did not. Within both subgroups, isolates from different Australian states grouped together. However, within the subgroup I clade, the NZ isolates formed three nested clades, and within the subgroup II clade, the NZ isolates formed two subclades – one with isolates from the North Island (Harrisville; HV18, and HV19), and one with isolates from the South Island (Sefton; SF2, and Nelson; RPE1, RPE3, and RPC3). Although one subgroup I nested clade contained isolates from both the North Island (Levin; WHG1-5) and the South Island (Sefton; SF1, and SF3; Figure 3.6), this could be an artefact of movement between lettuce growing areas. The spread of several plant viruses, including OFV (Kubo et al. 2009), is likely the result of human-mediated movement of infected material (Picard et al. 2017; Pagán 2018; Dietzgen et al. 2020). Indeed, the movement of lettuces infested with the *N. ribisnigri* aphid, a suspected vector for LNYV, contributed to the spread of the aphid to all NZ lettuce growing regions within 12 months of its discovery in the country (Stufkens and Teulon 2003; Walker et al. 2003). However, the phylogenetic subclades that included Sefton samples were still distinguished from the other subclade samples by a bifurcating branch, in both the subgroup I and subgroup II clades. These observations may be due to the sample size, and additional isolates from the Sefton region may reveal a fourth location specific subgroup I nested clade, or it could reveal nested clades within the subgroup II South Island subclade. The NZ isolates could be further analysed to assess their relationship to their geography and their evolutionary history using phylogeographic analyses (Buhay et al. 2009).

Phylogeography incorporates temporal and spatial information, with statistical models, to reconstruct phylogenies that allow evolutionary patterns, genetic diversity, population dynamics and molecular epidemiology to be evaluated (Picard et al. 2017). Several programmes are available for this type of analysis, including Nested Clade Phylogeographic Analysis (NCPA; Buhay et al. 2009), BEAST or BAyesian STructured coalescent Approximation (BASTA; Picard et al. 2017), Spatial Phylogenetics Reconstruction of Evolutionary Dynamics (SPREAD; Bielejec et al. 2011) and it's updated version; Spatial Phylogenetics Reconstruction of Evolutionary Dynamics using Data-Driven Documents (SpreaD3; Bielejec et al. 2016), as well as packages that operate within the statistical programme R, such as BioGeography with Bayesian Evolutionary Analysis in R Scripts (BioGeoBEARS; Matzke 2013). However, some approaches required analysis of multiple genes, accurate sampling, and thorough geographic and molecular clock data (Jridi et al. 2006; Panchal and Beaumont 2010; Picard et al. 2017). An approach appropriate to assess closely related, or within species, phylogeographic relationships is also required to assess the isolates from this study (Matzke 2013; Picard et al. 2017). As further data from each isolate was required, phylogeographic analysis was beyond the scope of the current study, but presents an interesting avenue for future research. Geographic relationships between virus isolates has been reported for other rhabdoviruses including CoRSV, EMDV, and TaVCV (Revill et al. 2005; Pappi et al. 2016; Ramalho et al. 2016), although specific phylogeographic analyses has yet to be reported for these viruses.

With the increase in lettuce crop losses in NZ, two hypotheses were suggested. Fletcher et al. (2019) hypothesised that the increase in NZ crop losses could be from an increase in subgroup II, or the arrival of a new LNYV isolate of Australian lineage. The subgroup diagnosis described in Chapter 2, showing only four of 18 isolates belonging to subgroup II, indicates subgroup II is not displaying a stronger presence in NZ at this time. In fact, the isolates from Levin (WHG1-5), a region reporting increased crop losses (Fletcher et al. 2019), belonged only to subgroup I. Growers from Mid Canterbury (Sefton) and Nelson also experienced increased crop losses, and the four subgroup II isolates collected in 2018 (SF2, RPC3, RPE1, and RPE3) originated from these regions. Although these isolates branched from an Australian lineage, they were still distinct from the Australian isolates and formed their own NZ grouping. Therefore, an isolate to support the hypothesis proposed by Fletcher et al. (2019) has yet to be identified. Although sample size could be a contributing factor, it is important to consider that the analysis performed in this study has focused on only the N gene. Analysis of the G gene and its protein, with its role in attachment to the aphid vector (Sections 1.2.1.1 and 1.2.3.1), may provide further information to identify an isolate of more recent or closer Australian origin in NZ. Studies of variants for the rhabdoviruses EMDV, OFV and RVCV included analysis of the G gene

along with the N gene, or even included analysis of all the viral genes (Pappi et al. 2016; Kondo et al. 2017; Jones et al. 2019). Analysis of multiple viral genes was beyond the scope of this current study. However, the inclusion of multiple genes, in particular the G gene, in future research has been indicated for several occasions described in this discussion.

From the analysis described in this Chapter, the current NZ LNYV subgroup population structure was updated, and novel taxonomic relationships were observed. Through the inclusion of 20 additional NZ LNYV isolates, including five subgroup II isolates, phylogenetic analysis topology and pairwise distance calculations have shown that subgroup I and subgroup II are distinct from each other. This indicates different evolutionary lineages, and possibly selective pressures, have been impacting the evolution of the two subgroups, and also their evolution in each country. It has been suggested that no new subgroup I isolates have been detected in Australia, due to subgroup II supplanting subgroup I (Callaghan and Dietzgen 2005; Higgins et al. 2016b). The pairwise distance calculations from this study suggested negative selective pressure was impacting on the subgroup II isolates, adding support to this hypothesis, although additional *in silico* analysis could be performed to confirm this further.

Phylogenetic topology from this study supports previous analyses, and indicated that subgroup I appeared earlier than subgroup II (Higgins et al. 2016b; Ajithkumar 2018). Within each subgroup, the NZ and Australian isolates were distinct, but taxonomic and evolutionary relationships were illustrated. For subgroup I, the NZ isolates showed shared evolutionary lineages by the formation of nested clades, and the Australian isolates a separate lineage. The geographic origin for this clade was unclear. For subgroup II, the NZ and Australian isolates showed some shared evolutionary history, with Australia indicated as the geographic origin for this clade. However, an isolate to support the hypothesis that increased crop losses in NZ were due to the arrival of a new variant from Australia, was not identified. With only four of 18 samples collected in 2018 identified as subgroup II, this would indicate that subgroup II has not been able to supplant subgroup I in NZ and is not exhibiting a stronger presence at this time. Although various further *in silico* analyses should be carried out in the future, the phylogenetic analysis described in this Chapter has generated valuable insight into the taxonomic and evolutionary relationships of NZ and Australian LNYV isolates. It has also provided information to assess hypotheses regarding recent increased crop losses in NZ, and the mechanism of displacement of subgroup I in Australia. The insights gained here and their relationship with the analyses carried out in Chapter 2 will be discussed further in Chapter 4.

Chapter 4

General Discussion

4.1 Final Discussion

Recently NZ lettuce growers have experienced increased crop losses in the lower North Island, Mid Canterbury, and Nelson regions of the South Island. Lettuce is a key economic crop (Section 1.1.1), and losses of up to 50 % were strongly associated with the cytorhabdovirus LNYV (Fletcher et al. 2019). Virus detection is important for maintaining, and increasing, the actual crop yield (Culliney 2014; Jeong et al. 2014; Savary et al. 2014; Sharma 2014). However, LNYV contains two viral subgroups, which cannot be distinguished by symptoms (Higgins et al. 2016b), and serological based methods are not appropriate for closely related viruses (Sections 1.2.4.1 and 2.1.2; Boonham et al. 2014; Jeong et al. 2014), therefore molecular based detection is required.

Several different primer sets have been developed and validated for identifying and diagnosing LNYV infection. However, these were not developed for distinguishing infection by each of the two subgroups (Sections 1.2.4.2.4 and 2.1). In 2018, Ajithkumar developed a diagnostic assay to distinguish between the two subgroups using RT-PCR-RFLP. However, for greater confidence, each sample needed to be tested twice – once with the subgroup I primers, and once with the subgroup II primers. The study presented here describes the development of an RT-qPCR-HRM assay to diagnose the two subgroups, the use of this assay to diagnose previously untyped NZ LNYV isolates (Chapter 2), and phylogenetic analysis of these isolates with other published LNYV isolates from NZ and Australia (Chapter 3). These analyses were performed to update the LNYV subgroup population structure in NZ, and further investigate the taxonomic and evolutionary relationships of the NZ and Australian LNYV isolates.

It has been suggested that the increase in NZ lettuce crop losses could be due to an increased presence of LNYV subgroup II in NZ, or a new strain of the virus from Australia (Fletcher et al. 2019). This study provided data to further assess these hypotheses. LNYV subgroup diagnosis using the two-step RT-qPCR-HRM analysis developed in this study identified 11 untyped isolates as subgroup I (WHG1-5, RPO1-4, RPE2, and SF3), four as subgroup II (RPC3, RPE1, RPE3, and SF2), and three as co-infected with both subgroups (SF1, RPC1, and RPC2). These results allowed our understanding of the NZ LNYV subgroup population to be updated, as well as diagnosis of apparent co-infection by both subgroups for the first time. The low number of identified subgroup II samples indicated that an increased presence of subgroup II was unlikely to be the cause of increased crop losses, as was suggested by Fletcher et al. (2019). Further, no subgroup II samples were identified from the lower North Island (Levin; WHG1-5), an area reported to be experiencing these crop losses (Fletcher et al. 2019). Although, subgroup II samples were identified from the two other regions experiencing crop losses, of the 13

samples from these regions there were only four identified as subgroup II. While the sample size may be too small to make definitive statements, the relatively small proportion of samples identified as subgroup II makes it unlikely that the cause of increased crop losses were due to LNYV subgroup II.

The identification of apparent co-infection of both LNYV subgroups in one host is important. The presence of co-infection could indicate greater virulence of subgroup II. Perhaps subgroup II is able to interact with insect or plant hosts in a more optimal manner (Higgins et al. 2016b), allowing subgroup II to infect a host plant already infected with subgroup I. The apparent extinction of subgroup I in Australia has been hypothesised to be due to subgroup II outcompeting subgroup I (Higgins et al. 2016b). The lack of subgroup I Australian isolates has made assessing this hypothesis in Australia more difficult but may be possible in NZ.

While RT-qPCR-HRM analysis identified apparent co-infection by the two LNYV subgroups, the phylogenetic analysis data did not appear to support this. Phylogenetic analysis of the entire N gene showed the amplified N gene of these isolates to be within the subgroup I clade (Chapter 3). This suggests only the subgroup I viral sequence was amplified during the one-step RT-PCR method used in Chapter 3. This was also noted when assessing the primer binding sites of Primer Set 3 against these N gene sequences (Chapter 2), which showed correlation with the subgroup I specific primers only (Section 2.3.4.7). If the subgroup I RNA was more abundant in the infected leaf tissue, it might be expected it would be preferentially amplified during the RT-PCR step over the less abundant subgroup II RNA. This could also be reflected in the major HRM melt peak identified as subgroup I that was seen with these samples after the two-step RT-qPCR-HRM analysis (Chapter 2). Further analysis is needed to measure the viral titres of these samples. In addition, future research to assess the viral titres of both subgroup positive control samples is warranted.

In plants infected with only subgroup II, the amount of viral RNA appeared to be higher than the amount amplified from the subgroup I infected plants. This might be due to inaccurate pipetting, greater qPCR efficiency (Nolan et al. 2013), or it might be suggestive of higher viral titres compared to the subgroup I positive controls (HV14, HV27). This in turn could be an indication of higher virulence of subgroup II, which may support the hypothesis that subgroup II was able to outcompete subgroup I in Australia due to its greater virulence (Higgins et al. 2016b). Measurement of the viral titres for both subgroup positive controls requires a quantitative analysis including a suitable reference gene (Bustin et al. 2009; Nolan et al. 2013) and was beyond the diagnostic scope of this study.

With molecular methods being the most suitable for LNYV subgroup detection (Boonham et al. 2014; Jeong et al. 2014), primer binding sites need to be highly virus subgroup specific. However, they also need to be flexible enough to detect the sequence variation observed within the two LNYV subgroups (Chapter 3; Callaghan and Dietzgen 2005; Higgins et al. 2016). This has also been noted for OFV and RVCV, where new viral variants were initially identified after symptoms and morphology matched, but no amplification was seen with primers originally designed for virus diagnosis (Jones et al. 2019; Roy et al. 2020). New primers containing degeneracies were designed in both cases to provide flexible detection of virus variants (Jones et al. 2019; Roy et al. 2020). Alternatively, in some situations, original primers may also detect variant virus strains. As was the case for TaVCV, where primers designed by Revill et al. (2005) from the Fijian TaVCV variant were able to amplify the appropriate region of the N gene from the first Hawaiian variant, even though the region showed only 84 % similarity between the two strains (Long et al. 2014). The primers developed in this study, for LNYV subgroup diagnosis using two-step RT-qPCR-HRM analysis (Chapter 2), were designed from LNYV N gene sequences from Australia and only one location in NZ – Harrisville, Auckland. This region has not reported recent increased lettuce crop losses (Fletcher et al. 2019), and were from one sampling in 2011 (Higgins et al. 2016b). However, the flexibility of the designed primers were still able to diagnose the LNYV subgroup of NZ isolates from other locations, collected seven years later in 2018 (Chapter 2).

In addition to the identification of co-infected samples, the primers designed in this study were flexible enough to distinguish LNYV sequence variation across sample technical replicates. This suggested the presence of a quasi-species nature for LNYV, as the primer binding sites were established to be subgroup specific (Chapter 2). Quasi-species population dynamics in studied plant viruses has indicated the host environment; including even different plant host cultivars, plays a key role in the population variation and diversity (Schneider and Roossinck 2001; Jo et al. 2017). Therefore, the quasi-species nature of LNYV may also be reflected by variation within the two subgroups that was identified by phylogenetics in Chapter 3 and in other N gene phylogenetic analyses (Callaghan and Dietzgen 2005; Higgins et al. 2016b).

The subgroup I clade formed three regional subclades within the NZ isolates, suggesting there was distinguishing variation between these isolates. Future research involving HTS, cloning and sequencing of isolates, or sequencing the cDNA of isolates, may further characterise the indicated quasi-species nature and distinguishing variation of LNYV (Benmansour et al. 1997; Schneider and Roossinck 2001; Ali and Roossinck 2017; Díaz-Martínez et al. 2018). Although, HRM analysis to differentiate variants of citrus HSVd (Loconsole et al. 2013), and genetic markers for PVY resistance (Nie et al. 2016) both reported correlation between HRM melt peak

groupings and formation of phylogenetic clades, this was not seen with LNYV. The sequence variation that distinguished the regional subclades of the LNYV isolates did not correlate with the normalised groupings generated by the HRM analysis. This illustrates other relationships in the sequences exist between these isolates and identifies a limitation when using diagnostic approaches unaided by other bioinformatics tools. Phylogenetic analyses can provide information on mutation rates, selective pressures, and molecular clock data (Lam et al. 2010), providing further information for molecular diagnostic designs (Bester et al. 2012; Jones et al. 2019).

Although such analyses were beyond the scope of this study, they reinforce the symbiotic relationship between diagnostics and phylogenetics. In addition, the phylogenetic analysis described in this study supports further bioinformatics research using these approaches into LNYV. The increased number of samples provided by this study increases the available taxonomic and sequence data for LNYV that could be used for future diagnostic primer design, updating the LNYV population structure, and characterising isolate relationships. These, together with molecular evolutionary data, such as substitution and mutation rates, could be used for assessing hypotheses regarding the epidemiology of LNYV and its connection to recent lettuce crop losses in NZ.

While country specific subclades were identified within each subgroup, variation within the NZ subgroup I isolates further identified regional subclades. Although the origin for the subgroup I clade was unclear, phylogenetic analysis of the translated protein sequences supported an Australian origin for the subgroup II clade. It has been suggested that LNYV is endemic to NZ and Australia (Randles and Carver 1971; Higgins et al. 2016b). Although sowthistle, the most important reservoir host for LNYV (Dietzgen et al. 2007), is not native to either country, other native plants have been identified as hosts (Randles and Carver 1971; Francki et al. 1989). New molecular clock or BEAST analysis was not performed in this study, but the ML phylogenetic trees reported here support relative emergence of each subgroup as calculated by Higgins et al. (2016), who suggested that subgroup I emerged earlier (150 years ago), with subgroup II emerging only 75 years ago. In Australia, once subgroup II emerged it appeared to outcompete subgroup I, – which then became extinct. During this time, subgroup I also emerged in NZ, and spread throughout the country. This is evidenced by the identification of subgroup I isolates at all regions sampled in both the North and South Islands. It would appear that subgroup II emerged more recently here, and from an Australian origin (Chapter 3). The recent emergence of subgroup II is indicated by the relatively low isolate numbers and, although subgroup II has been identified in both Islands, no isolates from the lower North Island (Levin) belonged to this subgroup. These low numbers would suggest subgroup II has not had time to supplant

subgroup I in NZ as it has in Australia. However, it is important to note that, for the first time, three isolates from the South Island were identified as subgroup co-infected, increasing the total number of subgroup II isolates.

4.1.1 Summary Conclusion

This study describes a sensitive, specific, and cost-effective two-step RT-qPCR-HRM assay used to diagnose the LNYV subgroup of previously untyped samples from both the North and South Islands of NZ. The development of this assay involved the testing of several different fluorescent dyes, concluding that the EvaGreen dye produced highly sensitive, reproducible and cost-effective results. The Ajithkumar Primers previously designed for use in RT-PCR-RFLP were unable to distinguish between the two subgroups in a two-step RT-qPCR-HRM assay, leading to the development of primers specific to this analysis. The developed primers were shown to be highly flexible and specific, identifying sequence variations across technical sample replicates, as well as three samples as co-infected with both subgroups. To date, this is the first instance of LNYV subgroup co-infection to be reported. Future research involving the measurement of viral titres, as well as sequencing of isolates or isolate cDNA, is recommended to elucidate further the virulence of the two subgroups and the possible quasi-species nature of LNYV identified in this study. Seven of the untyped samples were identified as subgroup II, including three co-infected samples, indicating an increased presence of subgroup II is unlikely to be the cause of increased crop losses in NZ.

Alternatively, it has been suggested that a new Australian LNYV strain has arrived here and is causing an increase in lettuce crop losses. To aid assessment of this, the unpublished isolates were sequenced and included in phylogenetic analysis with the N gene of other published NZ and Australian LNYV isolates. The entire N gene of the unpublished isolates was amplified by RT-PCR, the PCR product purified, and sequenced by Sanger sequencing. The nucleotide and translated protein sequences were used for ML tree construction with 1,000 bootstrap replicates including a total of 43 LNYV isolates. An isolate indicating a new Australian LNYV strain had arrived in NZ was not identified; however, other relationships were characterised. Within each subgroup clade the NZ and Australian isolates grouped by country. The subgroup I clade indicated a more stable population, and the NZ subgroup I isolates formed three well-supported subclades that appeared to be location specific. The subgroup II isolates showed a polytomy branching pattern with low bootstrap support, indicating rapid divergence and dispersal of the isolates. The two NZ subgroup II subclades branched from an Australian lineage, and the overall origin of this clade appeared to be Australian also. Further *in silico* analyses of these isolates is warranted to confirm the taxonomic and evolutionary

relationships that were characterised in this study. In particular, analysis of the G gene and its protein may provide useful information due to its interaction with the viral vector. It could aid understanding of selective pressures acting on LNYV sequences, could aid identification of an isolate of closer or more recent Australian origin, and could reveal possible indicators for the probable greater virulence of subgroup II. Nevertheless, this study provided data to assess several LNYV hypotheses, as well as updating our knowledge of the NZ LNYV subgroup population through the development of a two-step RT-qPCR-HRM diagnostic assay.

References

Adams I, Fox A. 2016. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis. In: Wang A, Zhou X, editors. Current research topics in plant virology. Switzerland: Pringer International Publishing. p. 323–335.

Ajithkumar P. 2018. Diagnosis and genome analysis of Lettuce necrotic yellows virus subgroups. Auckland University of Technology.

Ali A, Roossinck MJ. 2017. Analysis of quasispecies variation in single and mixed viral infection. Virus Evol. 3(2):1–7. doi:10.1093/ve/vex037.

Aliotta JM, Pelletier JJ, Ware JL, Moran LS, Benner JS, Kong H. 1996. Thermostable Bst DNA polymerase I lacks a 3′ → 5′ proofreading exonuclease activity. Genet Anal Biomol Eng. 12(5):185–195. doi:https://doi.org/10.1016/S1050-3862(96)80005-2. http://www.sciencedirect.com/science/article/pii/S1050386296800052.

Almeida RPP, Daane KM, Bell VA, Blaisdell GK, Cooper ML, Herrbach E, Pietersen G. 2013. Ecology and management of grapevine leafroll disease. Front Microbiol. 4(94):1–13. doi:10.3389/fmicb.2013.00094.

Aloisio M, Morelli M, Elicio V, Saldarelli P, Ura B, Bortot B, Severini GM, Minafra A. 2018. Detection of four regulated grapevine viruses in a qualitative, single tube real-time PCR with melting curve analysis. J Virol Methods. 257(July):42–47. doi:10.1016/j.jviromet.2018.04.008. https://doi.org/10.1016/j.jviromet.2018.04.008.

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. 2003. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol Evol. 19(10):535–544. doi:DOI 10.1016/j.tree.2004.07.021. http://linkinghub.elsevier.com/retrieve/pii/S0169534704002186%5Cnpapers://74e8b4ef -c919-4470-93db-e267bd9e4a22/Paper/p4577.

Applied Biosystems. 2009. A guide to high resolution melting (HRM) analysis. :20.

Arif M, Ochoa-Corona FM. 2013. Comparative assessment of 5′ A/T-rich overhang sequences with optimal and sub-optimal primers to increase PCR yields and sensitivity. Mol Biotechnol. 55(1):17–26. doi:10.1007/s12033-012-9617-5.

Baldauf SL. 2003. Phylogeny for the faint of heart: A tutorial. Trends Genet. 19(6):345–351.

doi:10.1016/S0168-9525(03)00112-4.

Bejerman N, Giolitti F, de Breuil S, Trucco V, Nome C, Lenardon S, Dietzgen RG. 2015. Complete genome sequence and integrated protein localization and interaction map for alfalfa dwarf virus, which combines properties of both cytoplasmic and nuclear plant rhabdoviruses. Virology. 483:275–283. doi:10.1016/j.virol.2015.05.001. http://dx.doi.org/10.1016/j.virol.2015.05.001.

Benmansour A, Basurco B, Monnier AF, Vende P, Winton JR, De Kinkelin P. 1997. Sequence variation of the glycoprotein gene identifies three distinct lineages within field isolates of viral haemorrhagic septicaemia virus, a fish rhabdovirus. J Gen Virol. 78(11):2837–2846. doi:10.1099/0022-1317-78-11-2837.

Beresford R, Mckay A. 2012. Climate change impacts on plant diseases affecting New Zealand horticulture.

Bester R, Jooste AEC, Maree HJ, Burger JT. 2012. Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate grapevine leafroll-associated associated virus 3 variant groups I, II, III and VI. Virol J. 9. doi:10.1186/1743-422X-9-219.

Bielejec F, Baele G, Vrancken B, Suchard M., Rambaut A, Lemey P. 2016. SpreaD3: interactive visualisation of spatiotemporal history and trait evolutionary processes. Mol Biol Evol. 33(8):2167–2169. https://beast.community/spread3.

Bielejec F, Rambaut A, Suchard M., Lemey P. 2011. SPREAD: Spatial Phylogenetic Reconstrustion of Evolutionary Dynamics. Bioinformatics. 27(20):2910–2912. https://beast.community/spread.

Bininda-Emonds O. 2009. An introduction to phylogenetic analysis. In: Lecture D1. Oldenburg: Institute for Biology and Environmental Science, Carl von Ossietzky Universitat Oldenburg. p. 283–299.

Bio-Rad. 2014. SsoFast ™ EvaGreen ® Supermix 200. Bio-Rad Lab.:1–2. http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10014647A.pdf.

Biology Online. 2005. Cap. Dictionary. [accessed 2019 Feb 13]. https://www.biology-online.org/dictionary/Cap.

Biotium. 2009. EvaGreen: The very best dye for qPCR and HRM. Biot Inc.(1).

Bispo PJM, Davoudi S, Sahm ML, Ren A, Miller J, Romano J, Sobrin L, Gilmore MS. 2018. Rapid detection and identification of uveitis pathogens by qualitative multiplex real-time PCR. Investig Ophthalmol Vis Sci. 59(1):582–589. doi:10.1167/iovs.17-22597. http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.17-22597%0Ahttp://iovs.arvojournals.org/data/journals/iovs/936670/i1552-5783-59-1-582.pdf.

Bluth MJ, Bluth MH. 2013. Molecular pathology techniques. Clin Lab Med. 33(4):753–772. doi:https://doi.org/10.1016/j.cll.2013.09.004. http://www.sciencedirect.com/science/article/pii/S0272271213000760.

Bluth MJ, Bluth MH. 2018. Molecular pathology techniques: Advances in 2018. Clin Lab Med. 38(2):215–236. doi:10.1016/j.cll.2018.03.004. https://doi.org/10.1016/j.cll.2018.03.004.

Boakye DB, Randles JW. 1974. Epidemiology of lettuce necrotic yellows virus in South Australia. III.* Virus transmission parameters, and vector feeding behaviour on host and non-host plants. Aust J Agric Res. 25(5):791–802. doi:10.1071/AR9740791.

Boonham N, Kreuze J, Winter S, van der Vlugt R, Bergervoet J, Tomlinson J, Mumford R. 2014. Methods in virus diagnostics: From ELISA to next generation sequencing. Virus Res. 186:20–31. doi:10.1016/j.virusres.2013.12.007. http://dx.doi.org/10.1016/j.virusres.2013.12.007.

Buhay JE, Crandall KA, Posada D. 2009. Nested clade phylogeographic analysis for conservation genetics. In: Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C, editors. Population Genetics for Animal Conservation. Cambridge: Cambridge University Press. p. 80–103.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, et al. 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 55(4):611–622. doi:10.1373/clinchem.2008.112797.

Callaghan B, Dietzgen RG. 2005. Nucleocapsid gene variability reveals two subgroups of lettuce necrotic yellows virus: Brief report. Arch Virol. 150(8):1661–1667. doi:10.1007/s00705-005-0528-7.

Carneiro FA, Stauffer F, Janeiro R De. 2005. Viral membrane fusion: is glycoprotein G of rhabdoviruses a representative of a new class of viral fusion proteins? Brazilian J Med Biol

Res. 38:813–823. doi:10.1590/S0100-879X2005000600002.

Chambers TC, Francki RIB. 1966. Localization and recovery of Lettuce necrotic yellows virus from xylem tissues of Nicotiana glutinosa. Virology. 29(4):673–676.

Cheng J, Jiang Y, Rao P, Wu H, Dong Q, Wu Z, Ding X, Guo J. 2013. Development of a single-tube multiplex real-time PCR for detection and identification of five pathogenic targets by using melting-curve analysis with EvaGreen. Arch Virol. 158(2):379–386. doi:10.1007/s00705-012-1493-6.

Chomič A, Winder L, Armstrong KF, Pearson MN, Hampton JG. 2011. Detection and discrimination of members of the family Luteoviridae by real-time PCR and SYBR ® GreenER^TM melting curve analysis. J Virol Methods. 171:46–52. doi:10.1016/j.jviromet.2010.09.028.

Chu PW., Francki RI. 1982. Detection of lettuce necrotic yellows virus by an enzyme-linked immunosorbent assay in plant hosts and the insect vector. Assoc Appl Biol. 100:149–156.

Clark MF, Adams AN. 1977. Characteristics of the microplate method of enzyme linked immunosorbent assay for the detection of plant viruses. J Gen Virol. 34(3):475–483. doi:10.1099/0022-1317-34-3-475.

Cleland J. 2013. World population growth; past, present and future. Environ Resour Econ. 55(4):543–554. doi:10.1007/s10640-013-9675-6.

Compans ER, Cooper G, Honjo AT, Koprowski H, Wagner H, Domingo ME, Domingo E. 2006. Current topics in microbiology and immunology quasispecies: Concept and implications for virology. Domingo E, editor. Germany: Springer Science+Business Media.

Coutts BA, Thomas-Carroll ML, Jones RAC. 2004. Analysing spatial patterns of spread of Lettuce necrotic yellows virus and Lettuce big-vein disease in lettuce field plantings. Ann Appl Biol. 145(3):339–343. doi:10.1111/j.1744-7348.2004.tb00391.x.

Cox KL, Devanarayan V, Kriaucinunas A, Manetta J, Montrose C, Sittampalam S. 2012. Immunoassay methods. In: Sittampalam G, Grossman A, Brimacombe K, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Caaveiro JM., Chung TD., et al., editors. Assay Guidance Manual [Internet]. 2019th ed. Bethesda (MD): Eli Lilly & Company and the National Centre for Advancing Translational Sciences. p. 1–38. https://www.ncbi.nlm.nih.gov/books/NBK53196/.

Crowley NC. 1967. Factors affecting the local lesion response of Nicotiana glutinosa to Lettuce necrotic yellows virus. Virology. 31(1):107–113. doi:10.1016/0042-6822(67)90013-X.

Culliney TW. 2014. Crop losses to arthropods. In: Pimentel D, Peshin R, editors. Integrated pest management. 3rd ed. Dordrecht: Springer Netherlands. p. 201–225.

Dal Santo P, Velthuis R. 2008. Strategic agrichemical review process.

Díaz-Martínez L, Brichette-Mieg I, Pineño-Ramos A, Domínguez-Huerta G, Grande-Pérez A. 2018. Lethal mutagenesis of an RNA plant virus via lethal defection. Sci Rep. 8(1):1–13. doi:10.1038/s41598-018-19829-6.

Diaz BM, Barrios L, Fereres A. 2012. Interplant movement and spatial distribution of alate and apterous morphs of Nasonovia ribisnigri (Homoptera: Aphididae) on lettuce. Bull Entomol Res. 102(4):406–414. doi:10.1017/S0007485311000745.

Diaz MH, Winchell JM. 2016. The evolution of advanced molecular diagnostics for the detection and characterization of Mycoplasma pneumoniae. Front Microbiol. 7(MAR):1–17. doi:10.3389/fmicb.2016.00232.

Dietzgen RG, Bejerman NE, Goodin MM, Higgins CM, Huot OB, Kondo H, Martin KM, Whitfield AE. 2020. Diversity and epidemiology of plant rhabdoviruses. Virus Res. 281(March). doi:10.1016/j.virusres.2020.197942.

Dietzgen RG, Callaghan B, Campbell PR. 2007. Biology and genomics of Lettuce necrotic yellows virus. Glob Sci Books.:85–92.

Dietzgen RG, Callaghan B, Wetzel T, Dale JL. 2006. Completion of the genome sequence of Lettuce necrotic yellows virus, type species of the genus Cytorhabdovirus. Virus Res. 118(1–2):16–22. doi:10.1016/j.virusres.2005.10.024.

Dietzgen RG, Francki RIB. 1988. Analysis of lettuce necrotic yellows virus structural proteins with monoclonal antibodies and concanavalin A. Virology. 166(2):486–494. doi:10.1016/0042-6822(88)90519-3.

Dietzgen RG, Francki RIB. 1990. Reducing agents interfere with the detection of lettuce necrotic yellows virus in infected plants by immunoblotting with monoclonal antibodies. J Virol Methods. 28(2):199–206. doi:10.1016/0166-0934(90)90035-E.

Dietzgen RG, Kondo H, Goodin MM, Kurath G, Vasilakis N. 2017. The family Rhabdoviridae:

mono- and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. Virus Res. 227:158–170. doi:10.1016/j.virusres.2016.10.010. http://dx.doi.org/10.1016/j.virusres.2016.10.010.

Dietzgen RG, Mann KS, Johnson KN. 2016. Plant virus-insect vector interactions: Current and potential future research directions. Viruses. 8(11):1–21. doi:10.3390/v8110303.

Dobhal S, Olson JD, Arif M, Garcia Suarez JA, Ochoa-Corona FM. 2016. A simplified strategy for sensitive detection of Rose rosette virus compatible with three RT-PCR chemistries. J Virol Methods. 232:47–56. doi:10.1016/j.jviromet.2016.01.013. http://dx.doi.org/10.1016/j.jviromet.2016.01.013.

Domingo E. 2010. Mechanisms of viral emergence. Vet Res. 41(38). doi:10.1051/vetres/2010010.

Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev. 76(2):159–216. doi:10.1128/mmbr.05023-11.

Downey N. 2014. Explaining multiple peaks in qPCR melt curve analysis. Integr DNA Technol. [accessed 2019 Jun 30]. https://sg.idtdna.com/pages/education/decoded/article/interpreting-melt-curves-an-indicator-not-a-diagnosis.

Du L, Shi W, Li X, Lan Y, Sun F, Fan Y, Zhou T, Zhou Y. 2019. A reverse-transcription loop-mediated isothermal amplification (RT-LAMP) assay for detecting the pathogen of maize rough dwarf disease in China. Australas Plant Pathol. 48(5):485–489. doi:10.1007/s13313-019-00649-9. https://doi.org/10.1007/s13313-019-00649-9.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797. doi:10.1093/nar/gkh340.

Egan AN, Schlueter J, Spooner DM. 2012. Applications of next-generation sequencing in plant biology. Am J Bot. 99(2):175–185. doi:10.3732/ajb.1200020.

Eigen M. 1993. The origin of genetic information: Viruses as models. Gene. 135(1–2):37–47. doi:10.1016/0378-1119(93)90047-7.

Eischeid AC. 2011. SYTO dyes and EvaGreen outperform SYBR Green in real-time PCR. BMC Res Notes. 4:2–6. doi:10.1186/1756-0500-4-263.

EzBioCloud T. 2018. Type strain and reference strain. Chun Lab Inc. [accessed 2019 Aug 1]. https://help.ezbiocloud.net/type-strain-and-reference-strain/.

Fang Y, Ramasamy RP. 2015. Current and prospective methods for plant disease detection. Biosensors. 5(3):537–561. doi:10.3390/bios5030537.

Fletcher J, Walker M, Davidson M, Paull S, Palmer A. 2019. Outdoor lettuce virus disease project 2016-2018 Year 1. A Plant & Food Research report prepared for: Vegetables New Zealand Incorporated. https://www.freshvegetables.co.nz/assets/Lettuce-virus-disease-project-Report-John-Flectcher-PFR-2016-2018-Year-1-FINAL.pdf.

Fletcher J, Zhang Y, Kean A., Davidson M. 2018. Outdoor lettuce virus disease project , year 2 report , 2018. A Plant & Food Research report prepared for: Vegetables New Zealand Incorporated.

Fletcher JD, France CM, Butler RC. 2005. Virus surveys of lettuce crops and management of lettuce big-vein disease in New Zealand. New Zeal Plant Prot. 58:239–244. doi:10.30843/nzpp.2005.58.4287.

Fox A, Mumford RA. 2017. Plant viruses and viroids in the United Kingdom: An analysis of first detections and novel discoveries from 1980 to 2014. Virus Res. 241(June):10–18. doi:10.1016/j.virusres.2017.06.029. http://dx.doi.org/10.1016/j.virusres.2017.06.029.

Francki RIB, Randles JW. 1970. Lettuce necrotic yellows virus. Assoc Appl Biol.(26). http://www.dpvweb.net/dpv/showadpv.php?dpvno=26.

Francki RIB, Randles JW, Dietzgen RG. 1989. Lettuce necrotic yellows virus. Assoc Appl Biol.(343). http://www.dpvweb.net/dpv/showdpv.php?dpvno=343.

Fry PR, Close RC, Procter CH, Sunde R. 1973. Lettuce necrotic yellows virus in New Zealand. New Zeal J Agric Res. 16(1):143–146. doi:10.1080/00288233.1973.10421173. http://www.tandfonline.com/doi/full/10.1080/00288233.1973.10421173.

Gachon C, Mingam A, Charrier B. 2004. Real-time PCR: What relevance to plant studies? J Exp Bot. 55(402):1445–1454. doi:10.1093/jxb/erh181.

Gallet R, Michalakis Y, Blanc S. 2018. Vector-transmission of plant viruses and constraints imposed by virus–vector interactions. Curr Opin Virol. 33:144–150. doi:10.1016/j.coviro.2018.08.005.

Ganley RJ, Bulman LS. 2016. Dutch elm disease in New Zealand: impacts from eradication and management programmes. Plant Pathol. 65(7):1047–1055. doi:10.1111/ppa.12527.

Gori A, Cerboneschi M, Tegli S. 2012. High-Resolution Melting analysis as a powerful tool to discriminate and genotype pseudomonas savastanoi pathovars and strains. PLoS One. 7(1):1–13. doi:10.1371/journal.pone.0030199.

Greer G, Saunders C. 2012. The costs of Psa-V to the New Zealand kiwifruit industry and the wider community. Wellington, New Zealand.

Gudnason H, Dufva M, Bang DD, Wolff A. 2007. Comparison of multiple DNA dyes for real-time PCR: Effects of dye concentration and sequence composition on DNA amplification and melting temperature. Nucleic Acids Res. 35(19). doi:10.1093/nar/gkm671.

Heim F, Lot H, Delecolle B, Bassler A, Krczal G, Wetzel T. 2008. Complete nucleotide sequence of a putative new cytorhabdovirus infecting lettuce. Arch Virol. 153(1):81–92. doi:10.1007/s00705-007-1071-5.

Higgins CM, Bejerman N, Li M, James AP, Dietzgen RG, Pearson MN, Revill PA, Harding RM. 2016a. Complete genome sequence of Colocasia bobone disease-associated virus, a putative cytorhabdovirus infecting taro. Arch Virol. 161(3):745–748. doi:10.1007/s00705-015-2713-7.

Higgins CM, Chang WL, Khan S, Tang J, Elliott C, Dietzgen RG. 2016b. Diversity and evolutionary history of Lettuce necrotic yellows virus in Australia and New Zealand. Arch Virol. 161(2):269–277. doi:10.1007/s00705-015-2626-5.

Higuchi R, Dollinger G, Walsh PS, Griffith R. 1992. Simultaneous amplification and detection of specific DNA sequences. Bio/Technology. 10(4):413–417. doi:10.1038/nbt0492-413. https://doi.org/10.1038/nbt0492-413.

Higuchi R, Fockler C, Dollinger G, Watson R. 1993. Kinetic PCR Analysis: Real-time monitoring of DNA amplification reactions. Bio/Technology. 11(9):1026–1030. doi:10.1038/nbt0993-1026. https://doi.org/10.1038/nbt0993-1026.

HortNZ HNZ. 2017. New Zealand domestic vegetable production : The growing story.

Hull A. 2019. Identification of reference genes and quantification of gene expression changes in Nicotiana glutinosa plants infected by subgroups I & II of Lettuce necrotic yellows virus. Auckland University of Technology.

Hulo C, de Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, Le Mercier P. 2011. ViralZone: A knowledge resource to understand virus diversity. Nucleic Acids Res. 39(Database issue):D576-82. doi:10.1093/nar/gkq901.

ICTV. 2020. International Committee on Taxonomy of Viruses. [accessed 2020 May 26]. https://talk.ictvonline.org/.

ICTV IC on T of V. 2019. Virus Taxonomy: The ICTV Report on Virus Classification and Taxon Nomenclature. Online Rep ICTV. [accessed 2020 Jun 2]. https://talk.ictvonline.org/ictv-reports/ictv_online_report.

Invitrogen TS. 2018. SuperScript IV One-Step RT-PCR System: A nonstop route to first-class RT-PCR results. :4.

Ito T, Suzaki K, Nakano M. 2013. Genetic characterization of novel putative rhabdovirus and dsRNA virus from Japanese persimmon. J Gen Virol. 94:1917–1921. doi:10.1099/vir.0.054445-0.

Jackson AO, Dietzgen RG, Goodin MM, Bragg JN, Deng M. 2005. Biology of plant rhabdoviruses. Annu Rev Phytopathol. 43(1):623–660. doi:10.1146/annurev.phyto.43.011205.141136. http://www.annualreviews.org/doi/10.1146/annurev.phyto.43.011205.141136.

Jansson L, Koliana M, Sidstedt M, Hedman J. 2017. Blending DNA binding dyes to improve detection in real-time PCR. Biotechnol Reports. 14:34–37. doi:10.1016/j.btre.2017.02.002. http://dx.doi.org/10.1016/j.btre.2017.02.002.

Jeong HJ, Kwon JK, Pandeya D, Hwang J, Hoang NH, Bae JH, Kang BC. 2012. A survey of natural and ethyl methane sulfonate-induced variations of eIF4E using high-resolution melting analysis in capsicum. Mol Breed. 29(2):349–360. doi:10.1007/s11032-011-9550-5.

Jeong J-J, Ju H-J, Noh J. 2014. A Review of Detection Methods for the Plant Viruses. Res Plant Dis. 20(3):173–181. doi:10.5423/RPD.2014.20.3.173. http://koreascience.or.kr/journal/view.jsp?kj=SMBRCU&py=2014&vnc=v20n3&sp=173.

Jo Y, Choi H, Kim SM, Kim SL, Lee BC, Cho WK. 2017. The pepper virome: Natural co-infection of diverse viruses and their quasispecies. BMC Genomics. 18(1):1–12. doi:10.1186/s12864-017-3838-8.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics. 8(3):275–282. doi:10.1093/bioinformatics/8.3.275.

https://doi.org/10.1093/bioinformatics/8.3.275.

Jones S, McGavin W, MacFarlane S. 2019. The complete sequences of two divergent variants of the rhabdovirus raspberry vein chlorosis virus and the design of improved primers for virus detection. Virus Res. 265(February):162–165. doi:10.1016/j.virusres.2019.03.004. https://doi.org/10.1016/j.virusres.2019.03.004.

Jordan IK, Sutter IV BA, McClure MA. 2000. Molecular evolution of the Paramyxoviridae and Rhabdoviridae multiple- protein-encoding P gene. Mol Biol Evol. 17(1):75–86. doi:10.1093/oxfordjournals.molbev.a026240.

Jridi C, Martin J-F, Marie-Jeanne V, Labonne G, Blanc S. 2006. Distinct viral populations differentiate and evolve independently in a single perennial host plant. J Virol. 80(5):2349–2357. doi:10.1128/jvi.80.5.2349-2357.2006.

Khalifa ME, Varsani A, Ganley ARD, Pearson MN. 2015. Comparison of Illumina de novo assembled and Sanger sequenced viral genomes: A case study for RNA viruses recovered from the plant pathogenic fungus Sclerotinia sclerotiorum. Virus Res. 219:51–57. doi:10.1016/j.virusres.2015.11.001. http://dx.doi.org/10.1016/j.virusres.2015.11.001.

Klerks MM, Lindner JL, Vaskova D, Spak J, Thompson JR, Jelkmann W, Schoen CD. 2004. Detection and tentative grouping of strawberry crinkle virus isolates. Eur J Plant Pathol. 110:45–52.

Koloniuk I, Fránová J, Sarkisova T, Přibylová J. 2018. Complete genome sequences of two divergent isolates of strawberry crinkle virus coinfecting a single strawberry plant. Arch Virol. 163(9):2539–2542. doi:10.1007/s00705-018-3860-4. https://doi.org/10.1007/s00705-018-3860-4.

Komorowska B, Fiore N, Zamorano A, Li R. 2014. Simultaneous detection of cherry necrotic rusty mottle virus and cherry green ring mottle virus using real-time PCR and high resolution melting analysis. Mol Cell Probes. 28(4):186–191. doi:10.1016/j.mcp.2014.03.002. http://dx.doi.org/10.1016/j.mcp.2014.03.002.

Kondo H, Hirota K, Maruyama K, Andika IB, Suzuki N. 2017. A possible occurrence of genome reassortment among bipartite rhabdoviruses. Virology. 508(March):18–25. doi:10.1016/j.virol.2017.04.027. http://dx.doi.org/10.1016/j.virol.2017.04.027.

Kubo KS, Stuart RM, Freitas-Astúa J, Antonioli-Luizon R, Locali-Fabris EC, Coletta-Filho HD,

MacHado MA, Kitajima EW. 2009. Evaluation of the genetic variability of orchid fleck virus by single-strand conformational polymorphism analysis and nucleotide sequencing of a fragment from the nucleocapsid gene. Arch Virol. 154(6):1009–1014. doi:10.1007/s00705-009-0395-8.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 33(7):1870–1874. doi:10.1093/molbev/msw054. https://doi.org/10.1093/molbev/msw054.

Kuzmin I V., Novella IS, Dietzgen RG, Padhi A, Rupprecht CE. 2009. The rhabdoviruses: Biodiversity, phylogenetics, and evolution. Infect Genet Evol. 9(4):541–553. doi:10.1016/j.meegid.2009.02.005.

Lam TT-Y, Hon C-C, Tang JW. 2010. Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infections. Crit Rev Clin Lab Sci. 47(1):5–49. doi:10.3109/10408361003633318. http://dx.doi.org/10.3109/10408361003633318.

Le DT, Vu NT. 2017. Progress of loop-mediated isothermal amplification technique in molecular diagnosis of plant diseases. Appl Biol Chem. 60(2):169–180. doi:10.1007/s13765-017-0267-y. https://doi.org/10.1007/s13765-017-0267-y.

Leeks A, Segredo-Otero EA, Sanjuán R, West SA. 2018. Beneficial coinfection can promote within-host viral diversity. Virus Evol. 4(2):1–12. doi:10.1093/ve/vey028.

Li YD, Chu ZZ, Liu XG, Jing HC, Liu YG, Hao DY. 2010. A cost-effective high-resolution melting approach using the EvaGreen dye for DNA polymorphism detection and genotyping in plants. J Integr Plant Biol. 52(12):1036–1042. doi:10.1111/j.1744-7909.2010.01001.x.

Loconsole G, Onelge N, Yokomi RK, Kubaa RA, Savino V, Saponari M. 2013. Rapid differentiation of citrus hop stunt viroid variants by real-time RT-PCR and high resolution melting analysis. Mol Cell Probes. 27:221–229. doi:10.1016/j.mcp.2013.07.003.

Long MH, Ayin C, Li R, Hu JS, Melzer MJ. 2014. First report of taro vein chlorosis virus infecting taro (Colocasia esculenta) in the United States. Plant Dis. 98(8):1160. doi:10.1094/PDIS-12-13-1277-PDN. https://doi.org/10.1094/PDIS-12-13-1277-PDN.

Lorenz TC. 2012. Polymerase chain reaction: Basic protocol plus troubleshooting and optimization strategies. J Vis Exp.(63):1–15. doi:10.3791/3998.

Luchi N, Pratesi N, Simi L, Pazzagli M, Capretti P, Scala A, Slippers B, Pinzani P. 2011. High-

resolution melting analysis: A new molecular approach for the early detection of Diplodia pinea in Austrian pine. Br Mycol Soc. 115(8):715–723. doi:10.1016/j.funbio.2011.05.005.

Luo A, Qiao H, Zhang Y, Shi W, Ho SY, Xu W, Zhang A, Zhu C. 2010. Performance of criteria for selecting evolutionary models in phylogenetics: A comprehensive study based on simulated datasets. BMC Evol Biol. 10(1). doi:10.1186/1471-2148-10-242.

Mann KS, Bejerman N, Johnson KN, Dietzgen RG. 2016a. Cytorhabdovirus P3 genes encode 30K-like cell-to-cell movement proteins. Virology. 489:20–33. doi:10.1016/j.virol.2015.11.028. http://dx.doi.org/10.1016/j.virol.2015.11.028.

Mann KS, Dietzgen RG. 2014. Plant rhabdoviruses: New insights and research needs in the interplay of negative-strand RNA viruses with plant and insect hosts. Arch Virol. 159(8):1889–1900. doi:10.1007/s00705-014-2029-z.

Mann KS, Johnson KN, Carroll BJ, Dietzgen RG. 2016b. Cytorhabdovirus P protein suppresses RISC-mediated cleavage and RNA silencing amplification in planta. Virology. 490:27–40. doi:10.1016/j.virol.2016.01.003. http://dx.doi.org/10.1016/j.virol.2016.01.003.

Mann KS, Johnson KN, Dietzgen RG. 2015. Cytorhabdovirus phosphoprotein shows RNA silencing suppressor activity in plants, but not in insect cells. Virology. 476:413–418. doi:10.1016/j.virol.2014.12.023. http://dx.doi.org/10.1016/j.virol.2014.12.023.

Mao F, Leung WY, Xin X. 2007. Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. BMC Biotechnol. 7:1–16. doi:10.1186/1472-6750-7-76.

Martin KM, Dietzgen RG, Wang R, Goodin MM. 2012. Lettuce necrotic yellows cytorhabdovirus protein localization and interaction map, and comparison with nucleorhabdoviruses. J Gen Virol. 93(4):906–914. doi:10.1099/vir.0.038034-0.

Martinez MA, Martrus G, Capel E, Parera M, Franco S, Nevot M. 2012. Quasispecies dynamics of RNA viruses. In: Witzany G, editor. Viruses: Essential Agents of Life. 1st ed. Dordrecht: Springer Science+Business Media. p. 21–42.

Martinez N, Ribeiro EA, Leyrat C, Tarbouriech N, Ruigrok RWH, Jamin M. 2013. Structure of the C-Terminal Domain of Lettuce Necrotic Yellows Virus Phosphoprotein. J Virol. 87(17):9569–9578. doi:10.1128/JVI.00999-13. http://jvi.asm.org/cgi/doi/10.1128/JVI.00999-13.

Matzke NJ. 2013. Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front Biogeogr. 5(4):242–248. doi:10.21425/F5FBG19694. https://escholarship.org/uc/item/44j7n141.

Maxam AM, Gilbert W. 1977. A new method for sequencing DNA. Proc Natl Acad Sci. 74(2):560 LP – 564. doi:10.1073/pnas.74.2.560. http://www.pnas.org/content/74/2/560.abstract.

McLean GD, Francki RIB. 1967. Purification of lettuce necrotic yellows virus by column chromatography on calcium phosphate gel. Virology. 31(4):585–591. doi:https://doi.org/10.1016/0042-6822(67)90186-9. http://www.sciencedirect.com/science/article/pii/0042682267901869.

McLean GD, Wolanski BS, Francki RIB. 1971. Serological analysis of lettuce necrotic yellows virus preparations by immunodiffusion. Virology. 43(2):480–487. doi:https://doi.org/10.1016/0042-6822(71)90319-9. http://www.sciencedirect.com/science/article/pii/0042682271903199.

Monis PT, Giglio S, Saint CP. 2005. Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. Anal Biochem. 340(1):24–34. doi:10.1016/j.ab.2005.01.046.

Montgomery JL, Sanford LN, Wittwer CT. 2010. High-resolution DNA melting analysis in clinical research and diagnostics. Expert Rev Mol Diagn. 10(2):219–240. doi:10.1586/erm.09.84.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harb Symp Quant Biol. 51(1):263–273. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0022965004&partnerID=40&md5=91dbd554ac1b490a9de4f722de249b9d.

Nair U. 2016. The real-time PCR digest. Bite Size Bio. [accessed 2020 Jun 3]. https://bitesizebio.com/29508/real-time-pcr-digest/.

National Center for Biotechnology Information N. 2004. GenBank. Bethesda Natl Libr Med. [accessed 2020 Feb 12]. www.ncbi.nlm.nih.gov/gene/.

NEB L. Isothermal amplification & strand displacement. New Engl BioLabs Inc. [accessed 2020 Jul 2]. https://international.neb.com/products/pcr-qpcr-and-amplification-

technologies/isothermal-amplification-and-strand-displacement/isothermal-amplification-and-strand-displacement.

Nie X, Sutherland D, Dickison V, Singh M, Murphy AM, De Koeyer D. 2016. Development and validation of high-resolution melting markers derived from Ry sto STS markers for high-throughput marker-assisted selection of potato carrying Ry sto. Phytopathology. 106(11):1366–1375. doi:10.1094/PHYTO-05-16-0204-R. http://apsjournals.apsnet.org/doi/10.1094/PHYTO-05-16-0204-R.

Nolan T, Huggett J, Sanchez E. 2013. Good practice guide for the application of quantitative PCR (qPCR).

Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, Hase T. 2000. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 28(12):E63–E63. doi:10.1093/nar/28.12.e63. https://pubmed.ncbi.nlm.nih.gov/10871386.

Novella IS, Presloid JB, Beech C, Wilke CO. 2013. Congruent evolution of fitness and genetic robustness in vesicular stomatitis virus. J Virol. 87(9):4923–4928. doi:10.1128/jvi.02796-12.

Osathanunkul M, Suwannapoom C, Osathanunkul K, Madesis P, De Boer H. 2016. Evaluation of DNA barcoding coupled high resolution melting for discrimination of closely related species in phytopharmaceuticals. Phytomedicine. 23(2):156–165. doi:10.1016/j.phymed.2015.11.018. http://dx.doi.org/10.1016/j.phymed.2015.11.018.

Pace NR, Sapp J, Goldenfeld N. 2012. Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. Proc Natl Acad Sci. 109(4):1011–1018. doi:10.1073/pnas.1109716109. http://www.pnas.org/cgi/doi/10.1073/pnas.1109716109.

Pagán I. 2018. The diversity, evolution and epidemiology of plant viruses: A phylogenetic view. Infect Genet Evol. 65(July):187–199. doi:10.1016/j.meegid.2018.07.033. https://doi.org/10.1016/j.meegid.2018.07.033.

Pallás V, Sánchez-Navarro JA, James D. 2018. Recent advances on the multiplex molecular detection of plant viruses and viroids. Front Microbiol. 9(SEP):1–11. doi:10.3389/fmicb.2018.02087.

De Palma A. 2018 Nov. Sanger sequencing: Still the gold standard? Lab Manag. https://www.labmanager.com/insights/2018/11/sanger-sequencing-still-the-gold-

standard-#.Xidwt2gzbIU.

Panchal M, Beaumont MA. 2010. Evaluating nested clade phylogeographic analysis under models of restricted gene flow. Syst Biol. 59(4):415–432. doi:10.1093/sysbio/syq022.

Pappi PG, Maliogka VI, Amoutzias GD, Katis NI. 2016. Genetic variation of Eggplant mottled dwarf virus from annual and perennial plant hosts. Arch Virol. 161(3):631–639. doi:10.1007/s00705-015-2705-7.

Pereira L, Gomes S, Barrias S, Fernandes JR, Martins-Lopes P. 2018. Applying high-resolution melting (HRM) technology to olive oil and wine authenticity. Food Res Int. 103(July 2017):170–181. doi:10.1016/j.foodres.2017.10.026. https://doi.org/10.1016/j.foodres.2017.10.026.

Petrzik K. 2012. Bioinformatic analysis of the L polymerase gene leads to discrimination of new rhabdoviruses. J Phytopathol. 160(7–8):377–381. doi:10.1111/j.1439-0434.2012.01919.x.

PFR TNZI for P and FRL. 2017. Eliminating disease from grapevines. Impact Case Study.:1–2. [accessed 2019 Aug 29]. https://www.plantandfood.co.nz/growingfutures/case-studies/testing-for-grapevine-leafroll.

PFR TNZI of P and FRL. 2019. Fresh Facts. Auckland.

Picard C, Dallot S, Brunker K, Berthier K, Roumagnac P, Soubeyrand S, Jacquot E, Thébaud G. 2017. Exploiting genetic information to trace plant virus dispersal in landscapes. Annu Rev Phytopathol. 55(1):139–160. doi:10.1146/annurev-phyto-080516-035616.

Pierce LR, Stepien CA. 2012. Evolution and biogeography of an emerging quasispecies: Diversity patterns of the fish viral hemorrhagic septicemia virus (VHSv). Mol Phylogenet Evol. 63(2):327–341. doi:10.1016/j.ympev.2011.12.024. http://dx.doi.org/10.1016/j.ympev.2011.12.024.

Plantwise. Lettuce necrotic yellows virus. [accessed 2017 Aug 3]. www.plantwise.org/KnowledgeBank/Datasheet.aspx?dsid=30283.

Promega Resources. PCR amplification. Promega Corp Website. [accessed 2019 Nov 2]. https://worldwide.promega.com/resources/guides/nucleic-acid-analysis/pcr-amplification/.

Puchta H, Sänger HL. 1989. Sequence analysis of minute amounts of viroid RNA using the

polymerase chain reaction (PCR). Arch Virol. 106(3):335–340. doi:10.1007/BF01313962. https://doi.org/10.1007/BF01313962.

Puli'uvea C. 2017. Molecular analysis of vanilla mosaic virus from the Cook Islands. Auckland University of Technology.

Quanta Bioscience. 2010. AccuMelt ^{TM} HRM SuperMix. :9–10.

Radvanszky J, Surovy M, Nagyova E, Minarik G, Kadasi L. 2015. Comparison of different DNA binding fluorescent dyes for applications of high-resolution melting analysis. Clin Biochem. 48:609–616. doi:10.1016/j.clinbiochem.2015.01.010. http://www.sciencedirect.com/science/article/pii/S0009912015000326.

Rajamäki ML, Lemmetty A, Laamanen J, Roininen E, Vishwakarma A, Streng J, Latvala S, Valkonen JPT. 2019. Small-RNA analysis of pre-basic mother plants and conserved accessions of plant genetic resources for the presence of viruses. PLoS One. 14(8):1–21. doi:10.1371/journal.pone.0220621.

Ramalho TO, Figueira AR, Wang R, Jones O, Harris LE, Goodin MM. 2016. Detection and survey of Coffee ringspot virus in Brazil. Arch Virol. 161(2):335–343. doi:10.1007/s00705-015-2663-0. https://doi.org/10.1007/s00705-015-2663-0.

Randles JW, Carver M. 1971. Epidemiology of Lettuce necrotic yellows virus in South Australia: II.* Distribution of virus, host plants, and vectors. Aust J Agric Res. 22(2):231–237. doi:10.1071/AR9710231.

Rao MS, Umamaheswari R, Priti K, Rajinikanth R, Grace GN, Kamalnath M, Prabu P, Kumar RM, Chaya MK, Vidyashree. 2016. Role of biopesticides in the management of nematodes and associated diseases in horticultural crops. In: Hakeem KR, Akhtar MS, Abdullah SNA, editors. Plant, Soil and Microbes: Volume 1: Implications in Crop Science. Cham: Springer International Publishing. p. 117–148. https://doi.org/10.1007/978-3-319-27455-3_7.

Redinbaugh M. G, Hogenhout S. A. 2005. Plant rhabdoviruses. In: Fu ZF, editor. The World of Rhabdoviruses. Vol. 18. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 143–163. https://doi.org/10.1007/3-540-27485-5_7.

Revill P, Trinh X, Dale J, Harding R. 2005. Taro vein chlorosis virus: Characterization and variability of a new nucleorhabdovirus. J Gen Virol. 86(2):491–499. doi:10.1099/vir.0.80591-0.

Roche. 2008. High Resolution Melting: Optimization strategies. Roche Diagnostics GmbH.(1).

Roche Diagnostics. 2012. LightCycler 96 Technical Specifications and User Manual.

Roy A, Stone AL, Melzer MJ, Shao J, Hartung JS, Mavrodieva V, Nakhla MK, Brlansky RH, Schneider WL. 2018. Complete nucleotide sequence of a novel hibiscus-infecting cilevirus from Florida and its relationship with closely associated cileviruses. Genome Announc. 6(4):e01521-17. doi:10.1128/genomeA.01521-17. http://mra.asm.org/content/6/4/e01521-17.abstract.

Roy A, Stone AL, Otero-Colina G, Wei G, Brlansky RH, Ochoa R, Bauchan G, Schneider WL, Nakhla MK, Hartung JS. 2020. Reassortment of genome segments creates stable lineages among strains of orchid fleck virus infecting citrus in Mexico. Phytopathology. 110(1):106–120. doi:10.1094/PHYTO-07-19-0253-FI.

Samarfard S, Bejerman NE, Dietzgen RG. 2018. Distribution and genetic variability of alfalfa dwarf virus, a cytorhabdovirus associated with alfalfa dwarf disease in Argentina. Virus Genes. 54(4):612–615. doi:10.1007/s11262-018-1563-2. http://dx.doi.org/10.1007/s11262-018-1563-2.

Samarfard S, McTaggart AR, Sharman M, Bejerman NE, Dietzgen RG. 2020. Viromes of ten alfalfa plants in Australia reveal diverse known viruses and a novel RNA virus. Pathogens. 9(3):1–24. doi:10.3390/pathogens9030214.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 74(12):5463–5467. doi:10.1073/pnas.74.12.5463.

Savary S, Ficke A, Hollier CA. 2014. Impacts of global change on crop production and food security. In: Freedman B, editor. Global environmental change. Dordrecht: Springer Netherlands. p. 379–387. https://doi.org/10.1007/978-94-007-5784-4_8.

Schaad NW, Frederick RD. 2002. Real-time PCR and its application for rapid plant disease diagnostics. Can J Plant Pathol. 24(3):250–258. doi:10.1080/07060660209507006.

Schneider WL, Roossinck MJ. 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. J Virol. 75(14):6566–6571. doi:10.1128/JVI.75.14.6566-6571.2001.

Selvarajan R, Balasubramanian V. 2016. Cutting-edge technologies for detection of plant viruses in vegetatively propagated crop plants. In: Gaur RK, Petrov NM, Patil BL, Stoyanova MI, editors. Plant viruses : evolution and management. 1st ed. Singapore:

Springer Nature. p. 312.

Sharma HC. 2014. Climate change effects on insects: Implications for crop protection and food security. J Crop Improv. 28(2):229–259. doi:10.1080/15427528.2014.881205. http://dx.doi.org/10.1080/15427528.2014.881205.

Sharma S, Kooner R, Arora R. 2017. Insect Pests and Crop Losses. In: Arora R, Sandhu S, editors. Breeding Insect Resistant Crops for Sustainable Agriculture. Singapore: Springer Singapore. p. 45–66. https://doi.org/10.1007/978-981-10-6056-4_2.

Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, Waterston RH. 2017. DNA sequencing at 40: Past, present and future. Nature. 550(7676). doi:10.1038/nature24286.

Shin DJ, Andini N, Hsieh K, Yang S, Wang T-H. 2019. Emerging Analytical Techniques for Rapid Pathogen Identification and Susceptibility Testing. Annu Rev Anal Chem. 12(1):annurev-anchem-061318-115529. doi:10.1146/annurev-anchem-061318-115529. https://www.annualreviews.org/doi/10.1146/annurev-anchem-061318-115529.

Sigma Aldrich. Sample purification and quality assessment. Merck. [accessed 2018 Oct 26]. https://www.sigmaaldrich.com/technical-documents/articles/biology/sample-purification-and-quality-assessment.html.

Simko I. 2016. High-resolution DNA melting analysis in plant research. Trends Plant Sci. 21(6):528–537. doi:10.1016/j.tplants.2016.01.004. http://dx.doi.org/10.1016/j.tplants.2016.01.004.

Simsek M, Adnan H. 2000. Effect of single mismatches at 3'-end of primers on polymerase chain reaction. Sultan Qaboos Univ Med J. 2(1):11–14.

Słomka M, Sobalska-Kwapis M, Wachulec M, Bartosz G, Strapagiel D. 2017. High resolution melting (HRM) for high-throughput genotyping-limitations and caveats in practical case studies. Int J Mol Sci. 18(11). doi:10.3390/ijms18112316.

Spiertz H. 2010. Food production, crops and sustainability: Restoring confidence in science and technology. Curr Opin Environ Sustain. 2(5–6):439–443. doi:10.1016/j.cosust.2010.10.006. https://pdf.sciencedirectassets.com/278668/1-s2.0-S1877343510X00057/1-s2.0-S1877343510001120/main.pdf?X-Amz-Security-Token=AgoJb3JpZ2luX2VjEO%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVh

c3QtMSJGMEQCIG7Dhqbu5A64giYgp6ca6aH4O0HcZ%2BDQ1MW3fyD4LpxGAiBPNb9caP2G.

StatsNZ SNZ. 2019. New Zealand's population grows 1.6 percent in June year. StatsNZ. [accessed 2019 Oct 2]. https://www.stats.govt.nz/news/new-zealands-population-grows-1-6-percent-in-june-year.

Stranneheim H, Lundeberg J. 2012. Stepping stones in DNA sequencing. Biotechnol J. 7(9):1063–1073. doi:10.1002/biot.201200153.

Stubbs LL, Grogan RG. 1963. Lettuce necrotic yellows virus. Nature. 4873:1229.

Stufkens MA., Teulon DA., Bulman S. 2002. Nasonovia ribis-nigri, a new aphid pest found on lettuces Lactuca sativa) and ribes spp. in Canterbury. [accessed 2017 Aug 9]. http://www.nzpps.org/journal/55/nzpp_554370.pdf.

Stufkens MAW, Teulon DAJ. 2003. Distribution, host range and flight pattern of the lettuce aphid in New Zealand. New Zeal Plant Prot. 56:27–33. doi:10.30843/nzpp.2003.56.6027.

Svec D, Tichopad A, Novosadova V, Pfaffl MW, Kubista M. 2015. How good is a PCR efficiency estimate : Recommendations for precise and robust qPCR efficiency assessments. Biomol Detect Quantif. 3:9–16. doi:10.1016/j.bdq.2015.01.005. http://dx.doi.org/10.1016/j.bdq.2015.01.005.

Tamura K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Mol Biol Evol. 9(4):678–687. doi:10.1093/oxfordjournals.molbev.a040752. https://doi.org/10.1093/oxfordjournals.molbev.a040752.

Tanno F, Nakatsu A, Toriyama S, Kojima M. 2000. Complete nucleotide sequence of Northern cereal mosaic virus and its genome organization. Arch Virol. 145(7):1373–1384. doi:10.1007/s007050070096.

Taylor S, Scott R, Kurtz R, Fisher C, Patel V, Bizouarn F. 2011. A practical guide to high resolution melt analysis genotyping. Bio-Rad Bull.:1–9.

ThermoFisher S. PCR Methods - Top Ten Strategies. ThermoFisher Sci. [accessed 2020a Jun 24]. https://www.thermofisher.com/nz/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-methods.html.

ThermoFisher S. PCR Setup - Six critical components to consider. ThermoFisher Sci. [accessed 2019b Oct 27]. https://www.thermofisher.com/nz/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-component-considerations.html#Mg.

ThermoFisher S. 2014. SYTO ® Green-Fluorescent Nucleic Acid Stains. :1–8. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/mp07572.pdf.

Thompson T, Batts W, Faisal M, Bowser P, Casey J, Phillips K, Garver K, Winton J, Kurath G. 2011. Emergence of viral hemorrhagic septicemia virus in the North American Great Lakes region is -associated with low viral genetic diversity. Dis Aquat Organ. 96(1):29–43. doi:10.3354/dao02362. http://www.int-res.com/abstracts/dao/v96/n1/p29-43/.

Thomson D, Dietzgen RG. 1995. Detection of DNA and RNA plant viruses by PCR and RT-PCR using a rapid virus release protocol without tissue homogenization. J Virol Methods. 54(2–3):85–95. doi:10.1016/0166-0934(95)00022-M.

Tirabassi R. How to quality control check your RNA samples. Bite Size Bio. [accessed 2018 Oct 26]. https://bitesizebio.com/13527/how-to-quality-control-check-your-rna-samples/.

Torrance L, Jones RAC. 1981. Recent developments in serological methods suited for use in routine testing for plant viruses. Plant Pathol. 30(1):1–24. doi:10.1111/j.1365-3059.1981.tb01218.x.

University of Waikato U. 2017. Three steps of PCR. Sci Learn Hub - Pokapū Akoranga Pūtaiao. [accessed 2020 Jun 5]. https://www.sciencelearn.org.nz/images/2983-three-steps-of-pcr%0A.

Varga A, James D. 2006. Real-time RT-PCR and SYBR Green I melting curve analysis for the identification of plum pox virus strains C, EA, and W: Effect of amplicon size, melt rate, and dye translocation. J Virol Methods. 132(1–2):146–153. doi:10.1016/j.jviromet.2005.10.004.

Veerakone S, Tang JZ, Ward LI, Liefting LW, Perez-Egusquiza Z, Lebas BSM, Delmiglio C, Fletcher JD, Guy PL. 2015. A review of the plant virus, viroid, liberibacter and phytoplasma records for New Zealand. Australas Plant Pathol. 44(5):463–514. doi:10.1007/s13313-015-0366-3.

Viruses are models for embracing diversity. 2018. Nat Microbiol. 3(4):389. doi:10.1038/s41564-018-0145-3.

Vunsh R, Rosner A, Stein A. 1990. The use of the polymerase chain reaction (PCR) for the detection of bean yellow mosaic virus in gladiolus. Ann Appl Biol. 117(3):561–569. doi:10.1111/j.1744-7348.1990.tb04822.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7348.1990.tb04822.x.

Waliullah S, Ling K-S, Cieniewicz EJ, Oliver JE, Ji P, Ali ME. 2020. Development of loop-mediated isothermal amplification assay for rapid detection of cucurbit leaf crumple virus. Int J Mol Sci. 21(5):1756. doi:10.3390/ijms21051756. https://pubmed.ncbi.nlm.nih.gov/32143404.

Walker G, Workman P, Stufkens M, Wright P, Fletcher J, Qureshi S, MacDonald F. 2003. IPM for outdoor lettuce.

Walker GT, Fraiser MS, Schram JL, Little MC, Nadeau JG, Malinowski DP. 1992a. Strand displacement amplification—an isothermal, in vitro DNA amplification technique. Nucleic Acids Res. 20(7):1691–1696. doi:10.1093/nar/20.7.1691. https://doi.org/10.1093/nar/20.7.1691.

Walker GT, Little MC, Nadeau JG, Shank DD. 1992b. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. Proc Natl Acad Sci. 89:392–396. doi:10.1073/PNAS.89.1.392. [accessed 2020 Jul 7]. https://www.pnas.org/content/89/1/392.

Walker PJ, Blasdell KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, Longdon B, Stone DM, Tesh RB, Tordo N, et al. 2018. ICTV virus taxonomy profile: Rhabdoviridae. J Gen Virol. 99(4):447–448. doi:10.1099/jgv.0.001020.

Walker PJ, Dietzgen RG, Joubert DA, Blasdell KR. 2011. Rhabdovirus accessory genes. Virus Res. 162(1–2):110–125. doi:10.1016/j.virusres.2011.09.004. http://dx.doi.org/10.1016/j.virusres.2011.09.004.

Wargo AR, Kurath G. 2012. Viral fitness: Definitions, measurement, and current insights. Curr Opin Virol. 2(5):538–545. doi:10.1016/j.coviro.2012.07.007. http://dx.doi.org/10.1016/j.coviro.2012.07.007.

Wetzel T, Dietzgen RG, Dale JL. 1994a. Genomic organization of Lettuce necrotic yellows rhabdovirus. Virology. 200:401–412.

Wetzel T, Dietzgen RG, Geering DW, Dale, J L. 1994b. Analysis of the nucleocapsid gene of lettuce necrotic yellows rhabdovirus. Virology. 202:1054–1057.

Whitfield AE, Huot OB, Martin KM, Kondo H, Dietzgen RG. 2018. Plant rhabdoviruses—their origins and vector interactions. Curr Opin Virol. 33:198–207. doi:10.1016/j.coviro.2018.11.002. https://linkinghub.elsevier.com/retrieve/pii/S1879625718300518.

Wieczorek D, Delauriere L, Schagat T. 2012. Methods of RNA quality assessment. Promega Corp Website.:1–14. [accessed 2018 Oct 26]. https://worldwide.promega.com/resources/pubhub/methods-of-rna-quality-assessment/.

Wong MH, Henderson J, Drenth A. 2013. Identification and differentiation of phyllosticta species causing freckle disease of banana using high resolution melting (HRM) analysis. Plant Pathol. 62(6):1285–1293. doi:10.1111/ppa.12056.

Yamagata Y, Yoshimura A, Anai T, Watanabe S. 2018. Selection criteria for SNP loci to maximize robustness of high-resolution melting analysis for plant breeding. Breed Sci. 68(4):488–498. doi:10.1270/jsbbs.18048.

Zhang Y, Xie Z, Fletcher JD, Wang Y, Wang R, Guo Z, He Y. 2020. Rapid and sensitive detection of lettuce necrotic yellows virus and cucumber mosaic virus infecting lettuce (Lactuca sativa) by reverse transcription loop-mediated isothermal amplification. Plant Pathol J. 36(1):76–86. doi:10.5423/PPJ.OA.12.2019.0298.