Viruses of imported Andean root and tuber crops and their plant health risk to the UK

Morgan Wodring

Supervisors: Dr. Adrian Fox
Prof. Neil Boonham
Dr. Kirsty McInnes
Dr. Fryni Drizou

Faculty of Science, Agriculture and Engineering
Newcastle University

This dissertation is submitted for the degree of
Doctor of Philosophy

January 2025

Abstract

Andean root and tuber crops (ARTCs) are a group of diverse crops belonging to a variety of families, that all originate from the Andean region of South America. From a UK perspective, these plants often enter the country for planting via internet trade, which is frequently a phytosanitary risk, and the CIP Genebank has identified developing tools for phytosanitary diagnostics of ARTCs as an area of concern.

This PhD followed on from a High Throughput Sequencing (HTS) study of ulluco (*Ullucus tuberosus*) plants grown in the UK. Its aim was to determine the virome of ARTCs imported through internet trade and determine the potential risk of this commodity to UK plant health. This project used HTS to study the viruses present in oca (*Oxalis tuberosa*) plants purchased from several European countries via eBay as well as oca, ulluco, yacón (*Smallanthus sonchifolius*) and mashua (*Tropaeolum tuberosum*) plants from a commercial supplier in Ireland. An Illumina MiSeq was used to sequence two bulks of these plants uncovered multiple novel virus candidates from different genera in three ARTCs. Following *in silico* molecular characterisation based on genome sequence, RT-PCR and qRT-PCR assays were designed to these viruses. A survey was conducted on further oca plants purchased from online sellers in Europe. Three novel viruses (OAV1, OCV1 and OVX) were detected in plants from several European countries, including samples of naturalised *Oxalis* species growing in the UK.

In addition, the quarantine virus potato yellowing virus (PYV), known to infect yacón, was found again in European plants sourced from European nurseries. The sequence of the viruses detected in these plants was closely related but not identical to sequences from a 2019 study of PYV infecting yacón, confirming that PYV continues to circulate in yacón grown in Europe and that yacón presents a plant health risk in Europe.

I'd like to thank some of the people who watched me totter around the labs and lent a hand: First, my amazing supervisors: Neil Boonham, Fryni Drizou, Adrian Fox and Kirsty McInnes, and also Lisa Ward, who supervised me at the beginning of this PhD and helped me find my feet. I couldn't have asked for more helpful, learned and kind people to supervise me. The gratitude I feel towards you for being there when I needed you wouldn't fit into an acknowledgments page, but just know that you went above and beyond for me and I think you're all wonderful people. Thanks also to the Virology team at Fera - all of you, but particularly Aimee Fowkes, Kinda Alraiss, Kiera Chisnall, Anna Skelton, Steven Forde, George Lawrence and Mary Mynett for helping with the work, and especially for being such good company! While we're talking about Fera, I also owe thanks to Ian Adams and Sam McGreig. Thanks for your enthusiasm, for your patience, and for your help. What I know about bioinformatics, I know from you.
Thanks to Dr. Inés Vazquez-Iglesias - the other 'science sibling'. You've been an immeasurable help over these years, probably more than you know. Here's to finding more allexiviruses in geese in future! Lastly to the people in my personal life who make it all worth it. Sukie Baker - my brain-twin who was only 16 days late. The rest of the Four Humours: Mara Reid and Argent Kolb - the other two brain cells - and of course Meiren Park, everyone's favourite turtle - 사랑해.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 80,000 words excluding appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Morgan Wodring
January 2025

Acknowledgements

This work would not have been possible without my supervisors (Adrian Fox, Neil Boonham, Kirsty McInnes and Fryni Drizou). In addition, high throughput sequencing datasets from prior to this PhD was utilised. Where a dataset originates from this 2019 material, this is labelled clearly in the text. In particular the work of Ian Adams of Fera Science Ltd. in categorising the contents of these datasets is acknowledged and thanked.
Work was included from Roy Macarthur of Fera Science Ltd. who provided statistical analysis and advice. His specific contributions of equations and R scripts are noted in the text.
I thank the Virology, NGS and MTU teams at Fera Science Ltd. for providing support, advice, and training. In particular, I collaborated with Chris Conyers on the RCA work of this thesis; with Sam McGreig and Ian Adams in bioinformatic analysis and with Aimee Fowkes and Kinda Alraiss in testing yacón for PYV using PCR and qPCR.

Acronyms

ARTCs Andean Root and Tuber Crops

cDNA Complimentary DNA

CIP The International Potato Centre (Centro Internacional de la Papa)

cnPCR Conventional PCR(used for Reverse-Transcription Conventional PCR here)

CP Coat Protein

CPM Commission on Phytosanitary Measures

DEFRA Department for Environment Food and Rural Affairs

DNA Deoxyribonucleic Acid

dsRNA Double-Stranded RNA

ELISA Enzyme-linked Immunosorbent Assay

EPPO European and Mediterranean Plant Protection Organization

EU European Union

EVE Endogenous Viral Element

GATT General Agreement on Tariffs and Trade

HTS High Throughput Sequencing

ICTV International Committee on Taxonomy of Viruses

IPPC International Plant Protection Convention

ISPM International Standard for Phytosanitary Measures

LCA Last Common Ancestor

MGW Molecular Grade Water

MP Movement Protein

NABP Nucleic Acid Binding Protein

NPPO National Plant Protection Organisation

ORF Open Reading Frame

PCR Polymerase Chain Reaction

PRA Pest Risk Assessment

PRONC Phytosanitary Risks of Newly Introduced Crops

qRT-PCR Quantitative Reverse-Transcription PCR

RACE Rapid Amplification of cDNA Ends

RCA Rolling Circle Amplification

RdRp RNA-Dependent RNA Polymerase

RHS Royal Horticultural Society

RNA Ribonucleic Acid

rRNA Ribosomal RNA

RT-PCR Reverse-Transcription PCR

siRNA Short Interfering RNA

SPS Agreement Agreement on the Application of Sanitary and Phytosanitary Measures

SRA Short Read Archive

TBP Triple Gene Block

totRNA Total RNA

UK United Kingdom

USA United States of America

UTR Untranslated Region

VANA Virion Associated Nucleic Acids

WTO World Trade Organisation

APLV	Andean potato latent virus	**OOV1**	Oca ophiovirus 1

Continued on next page

(Continued)

ArMV	Arabis mosaic virus
AVB	Arracacha virus B
BBWV2	Broad bean wilt virus 2
BLMoV	Blueberry leaf mottle virus
BLSV	Blueberry latent spherical virus
BRSV	Beet ringspot virus
CaMV	Cauliflower mosaic virus
CarMV	Carnation mottle virus
CLRV	Cherry leaf roll virus
CMV	Cucumber mosaic virus
DvEPRS	Dahlia variabilis endogenous pararetroviral sequence
GBLV	Grapevine Bulgarian latent virus
GFLV	Grapevine fanleaf virus
MVX	Mint virus X
MVY	Mashua virus Y
OAV1	Oca allexivirus 1
OcaV1	Oca caulimovirus 1
OCV1	Oca capulavirus 1
ONV1	Oca nepovirus 1
ONV2	Oca nepovirus 2
OPMaV1	Opium poppy mosaic associated virus
OVX	Oca virus X
PapMV	Papaya mosaic virus
PBRSV	Potato black ringspot virus
PLRV	Potato leafroll virus
PvaMV-8	Plasmopara viticola lesion associated mononegaambi virus 8
PVT	Potato virus T
PVX	Potato virus X
PVY	Potato virus Y
PYV	Potato yellowing virus
TBRV	Tomato black ring virus
TMV	Tobacco mosaic virus
ToRSV	Tomato ringspot virus
TropMV	Tropaeolum mosaic virus
TropPolV1	Tropaeolum polerovirus 1
TropPV3	Tropaeolum potyvirus 3
TRSV	Tobacco ringspot virus
TuMV	Turnip mosaic virus
UlPolV1	Ullucus polerovirus 1
UMMV	Ullucus mild mottle virus

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Throughout human history, plants - their seeds and other propagules - have been traded for food, medicine and aesthetic purposes. Trade in plants and their products is documented to be ancient. Garlic (*Allium sativum*) is a plant with a complex origins, with at least one group originating in middle Asia and being traded in ancient times as far as North China. This transfer is attested in ancient Chinese literature and occurred via trade from Iran or Turkeminstan no more than 1500 years ago (Maaß and Klaas, 1995). With trade of plants came an unintentional trade in their associated pests and pathogens. Plant diseases have been recorded in historical texts: the ethnographic Florentine Codex references smut and ear rot of corn as it was known to the Nahua farmers at the time (Orlob, 1971). This text was written by a Spanish writer during the 'Columbian Exchange'; the exchange of crops, disease, people and culture between the 'Old World' and the 'New World' following Columbus' arrival in the Americas in 1492 (Nunn and Qian, 2010). The European invaders extracted 'New World' germplasm material such as potato, chilli, and tomato to be grown overseas, and tribute was often taken by Iberian settlers in the form of local agricultural products rather than introduced species. Europeans in return brought barley, wheat and alliums like onion, garlic and leek which were adopted alongside traditional crops and soon found at markets in Tenochtitlán (Hernándo Bermejo and León, 1994). The Europeans also brought communicable diseases such as smallpox, which was more destructive to the indigenous people of the Americas than the Black Death was to Europeans (Nunn and Qian, 2010).

Exchange of communicable diseases of plants has also had a profound effect on human life. Famously, the Irish Potato Famine (*An Gorta Mór*) of 1845-51 led to as many as one and a half million excess deaths and many more emigrations, the scale of which is still not certain (Kelleher, 2009). The failure of the potato harvest was caused by *Phytophthora infestans*. The blight was not new to Europe, however, and the late blight along with other factors led to a failure of the potato harvest which may well have shaped the revolutionary period of European history, where hunger and disease led to social upheaval (Zadoks, 2008). The famine caused by the failure of the potato harvest in Ireland had intertwined societal causes stemming back decades, remaining a hotly contested and politicised subject. Social conditions led the Irish population to be vulnerable to the failure of the potato crop and lack access to other foods; scholars debate the extent to which English colonial rule may have been more lethal than the blight itself (Fraser, 2003, Kelleher, 2009). The murrain played an important part in European history and, with mass emigrations of hungry people, the world at large.

The large scale European famine caused in part by the late blight may be responsible for the global spread of potato virus Y isolates in the 19th century, due to concentrated breeding efforts that predated virus testing (Fuentes et al., 2019). Such an association between humankind and the viruses of plants we value is evident also in potato virus X, whose genetic variation maps to trade patterns large and small over the last several thousand years (Fuentes et al., 2021). These issues are relevant still today as plant disease, food security and trade are complex, interlinked subject with a pivotal human element.
While viruses of cash crops (e.g. tomato) are considered most economically important (Rybicki, 2015), those damaging staple crops such as cassava, maize and sweet potato in food-insecure regions are presently seen as the most dangerous to human life (Jones and Naidu, 2019). Cassava mosaic virus disease has been present in Africa since at least the late 19th century, but became epidemic in the 1990s, contributing to widespread food insecurity during a drought (Vurro et al., 2010). Yield losses can be as high as 50-70% in infected plants (Muimba-Kankolongo et al., 1997).
Some viruses emerge from native flora onto introduced hosts (Jones, 2020). Others disseminate due to the practice of multinationals growing plants for seed in countries with climates amenable to it, then exporting the infected seed (Jones, 2021). In general, seeds present the greatest threat to plant health when traded internationally, as planting infected seeds leads to the earliest possible infection of a plant with a virus (Vurro et al., 2010; Jones, 2006). As well as seeds, fresh propagules such as garlic bulbs imported for vegetative propagation are another source of virus spread (Wylie et al., 2014). Human mediated, long distance movement of plants and their viruses, and the risk that may pose, is the focus of the present thesis.

1.1. Trade and Virus Control

Links exist between trade volumes and biological invasions of all kinds, explained in part by the principle of propagule pressure (Meyerson and Mooney, 2007). As global trade increases, it is likely that biological invasions will continue to do so as well (Levine and D'Antonio, 2003). Invasion risk is also influenced by the biology of the invasive organism, though this can be less important than the simple availability and frequency of propagules for some taxa, due to stochastic effects (Simberloff, 2009).
International agreements have been put in place to regulate this balance of free trade and biosecurity. In 1947 the General Agreement on Tariffs and Trade (GATT) was formed with 23 members, later becoming the World Trade Organisation (WTO) in 1995 (Crowley, 2003). At its formation the WTO established the SPS Agreement (Agreement on the Application of Sanitary and Phytosanitary Measures), including the plant health portion of the agreement, the IPPC (International Plant Protection Convention) and its governing body, the CPM (Commission on Phytosanitary Measures). The stated aim of these agreements was to reduce biological invasions of all kinds, especially those as a result of international movement of goods (Kahn and Pelgrim, 2010). The WTO also arbitrates cases where a member state believes an SPS measure imposed

by another member is overly restrictive or protectionist in nature (Crowley, 2003). In order, negotiations were made to facilitate free, then fair, then safe trade (Whattam et al., 2014). Countries may take a precautionary approach (excluding all potential threats) or a known threats approach, where only pests on lists of threats are managed (Webber, 2010). The manner in which a governing body works with scientific uncertainty tends to influence the choice of a precautionary vs. known threats approach, with some countries viewing the precautionary principle as disguised protectionism (Hasbun, 2009). Conversely, an over-reliance on lists of known pests may allow unrecognised plant pests to enter and establish in a country (Brasier, 2005). Failure to manage plant health threats can result not just in losses due to injured plants, but opportunity costs due to reduced market access (Hulme, 2014).

The biosecurity options available to an importing country begin before the border with said pest risk analysis and certification, extend to the border with inspections and post-entry quarantine, and encompass post-border measures such as surveys and outbreak monitoring (Sequeira and Griffin, 2014). Generally, adapting to a plant pest outbreak is more costly than preventing one, and biosecurity's main aim is to prevent this additional cost through exclusion (Rodoni, 2009). Under current IPPC regulations, phytosanitary measures imposed by member states must have a scientific basis, justified via a risk assessment with accompanying risk management options (Whattam et al., 2014). Biosecurity is covered under IPPC rules by a series of international standards for phytosanitary measures (ISPMs). The ISPMs governing quarantine and regulated non-quarantine pests are 2, 11 and 21 (IPPC Secretariat, 2019, IPPC Secretariat, 2013, IPPC Secretariat, 2021). Surveillance is undertaken by the national plant protection organisation (NPPO) of the relevant countries (IPPC Secretariat, 2013). For the UK and most of Europe this is the European and Mediterranean Plant Protection Organization (EPPO).

In the EPPO region, quarantine pests are known as A1 pathogens and are absent from the EPPO region entirely, such as potato virus T, while A2 pathogens are regulated non-quarantine pests associated with plants for planting, and are present to some extent within the EPPO region, such as potato spindle tuber viroid (EPPO, 2023a; EPPO, 2023b; EPPO, 2023d). Phytosanitary certification is covered by ISPMs 7 and 12, and is used to assert that plant products exported by one country meet the import requirements of another (Whattam et al., 2014). Inspection agencies must work within their capacity to apply inspections in the most cost-effective way in order to minimise failure to detect infected consignments which may then enter the country (Rodoni, 2009). Additionally, countries are obligated to look for pests within their territory, i.e. to demonstrate that a pest is 'known not to occur' rather than 'not known to occur' (Van Halteren, 2000). In particular, free trade agreements between states, such as the EU, means that a phytosanitary practice is only as strong as its weakest link, making good biosecurity practice essential (Brasier, 2005). These purposes rely on diagnostic methods which are suitable to the task and thus reduce disputes between trade partners (Whattam et al., 2014). Thus understanding pathogens, their risk, and how to reliably identify them is the basis for biosecurity in any country both legally (fulfilling obligations) and practically (safeguarding plant health) (Trontin et al., 2021, Whattam et al., 2014).

1.2. e-Commerce as an Invasion Pathway for Plant Pests

The risks that e-Commerce in particular poses remain a topic of investigation by several governing bodies, including the IPPC (IPPC Secretariat, 2023a). In 2021, 73% of consumers on the European continent shopped online, including 91% of people from the UK, 36% of which was cross-border (Lone et al., 2021). Purchases of plants for planting from the internet have grown increasingly popular among small growers with gardens and allotments; in some cases this trade can be illegal. For example, unregulated imports of Goji plants from China were detected by the UK Plant Health and Seed Inspectorate (PHSI) in 2007-2008 being imported to the island of Guernsey, totalling 45,000 plants supplied to 19,000 customers (Giltrap et al., 2009).

A number of biological invasions have been linked to online trade in Aoteoroa-New Zealand, for which the best management was deemed to be public awareness and communication with online vendors (Derraik and Phillips, 2010). In terms of plant health, Aoteoroa-New Zealand's NPPO have faced a greater number of phytosanitary investigations as the result of breaches in recent years, most probably due to e-commerce and its ability to bypass import safeguards (Pearson and Hill, 2014). For example, strawberries are being imported for planting without undergoing the mandatory 6 month quarantine, and Pearson et al. (2014) also noted incorrect declaration of planting material as another, safer commodity. In Australia, curcubit seeds purchased over a six month period from abroad through the internet were found to contain viruses in 23/31 seed lots tested, some of which are regulated as quarantine pests in Australia (Constable et al., 2021). The availability of cheap, illegally imported plants through mail order may undermine the effectiveness of voluntary plant health certification programmes (Green et al. 2021).

Internet trading has allowed a greater variety of sellers to enter the market, including hobbyists and collectors, and has opened up a market of informal exchange, such as via Facebook and other social media sites (Invasive Species Advisory Committee, 2012). Management attempts have already failed for numerous plant pathogens along conventional pathways (Evans et al., 2010), and so e-commerce presents a great risk to plant health. A survey of international horticulturists revealed that a number of growers view plant health measures, particularly in Europe, to be insufficient when it comes to the 'grey sector' of tourists, car boot sales and international sales via courier (Yeomans et al., 2021). A study on internet trade found that invasive plants are both commonly available for sale and difficult to monitor (Humair et al., 2015). Generally, e-commerce is broken down into business-to-business, business-to-consumer, and person-to-person, in roughly ascending difficulty of regulatory significance (IPPC Secretariat, 2023a).

E-commerce can challenge traditional means of pre-border entry control. In particular, the following properties of e-commerce are most salient here:

1. Shipping is frequently via private couriers in small parcels. Private couriers often have no legal right to open packages. If they do, there are instances where a company may be hesitant to do so in case of client dissatisfaction. In addition, these couriers and their checks often lack the expertise typical of conventional customs agents: those involved in

identifying biosecurity threats, and how to deal with them (IPPC, 2012). In some cases, couriers may not be aware that the item they are shipping is biological at all (Invasive Species Advisory Committee, 2012).

2. Small packages arrive via ports and quickly disseminate to many final destinations, rather than the traditional large shipments that have one intermediate destination before the point of sale (Giltrap et al., 2009). Should a commodity have a high incidence of infection, this can make outbreaks difficult to eradicate should they occur. ISPM 11 considers the likelihood of a commodity being sent to "a few or many destination points in the PRA area" as part of its Pest Risk Assessment process (IPPC Secretariat, 2013).

3. e-commerce websites such as eBay and Facebook marketplace rely on legal conduct, as well as correct labelling from their individual sellers. Ignorance of international restrictions on plant products from both buyers and sellers, or intentional misdirection of authorities through incorrect information on customs forms, both lead to pest spread (IPPC Secretariat, 2023a). In some cases the internet can provide information on how to intentionally bypass customs through falsified declarations and trans-shipments (Invasive Species Advisory Committee, 2012). A study in Germany found that 39% of plants they ordered via the internet were intentionally mislabelled to evade customs inspections (Kaminski et al., 2012).

1.3. Transmission of Viruses

Plant viruses spread by several means, typically, but not solely, requiring a vector, the biology of which influences the epidemiology of the virus (Hull, 2001). A small number of viruses, characteristically members of the genus *Tobamovirus*, spread by mechanical contact when leaves of an infected plant touch an uninfected plant (Ilyas, 2022). Some viruses are vectored by the feeding of nematodes upon plant tissues such as roots and tubers, these include some nepoviruses; the 'ne' in the name of the group of virus refers to the nematode transmission (Fuchs et al. 2017). Viruses may be spread by herbivorous insects such as aphids when they feed on plants and viruses spread in this manner include the phloem-limited poleroviruses (persistent spread), caulimoviruses (semi-persistent spread) and non-persistent spread (potyviruses), determined by the region of the vector with which the virus interacts and how long a virus can be transmitted by the vector after virus acquisition (Miller, 1999; Gibbs et al., 2020; Hull, 2001). Eriophyid mites may vector some viruses, such as particular members of the genus *Allexivirus* (Mansouri and Ryšánek, 2021). Pollen transmission is another method of spread with epidemiological consequences; viruses which spread this way include cherry leaf roll virus, a nepovirus (Rowhani et al., 2017).

1.4. Andean Root and Tuber Crops

As a case study, the Andean Root and Tuber Crops (ARTCs) were chosen for this thesis. A large part of the reason is that they are commonly available from European sellers on websites such as eBay, shipped directly to one's home.

Generally for all plant pathogens, pre-border prevention of entry in the first place is considered the ideal (Sequeira and Griffin, 2014). Viruses are of particular concern to plant health professionals, and farmers, because there are no direct chemical cure, meaning that control must be through resistant strains and prophylaxis (Rubio et al., 2020), with limited success being found in inoculating plants with attenuated virus, as in humans (Kobayashi et al., 2014). Use of pesticides to control invertebrate vectors has hefty environmental costs and may even be counter-productive in an epidemic situation (Jones, 2006). Cultural control methods include fallowing or rotation with virus-resistant crops and can allow a field to recover, for example with some nematode-transmitted nepoviruses (Harrison, 1977). This approach is dependent on vector types and reservoir species. Early detection is important, then, to prevent and mitigate disease outbreaks, particularly in cases where infections are mixed and information is scarce (Vivek et al., 2020).

By understanding something of the virome of ARTCs traded via the internet, additional information on the pathway can be gleaned. Because of the European source of the ARTC samples sequenced during this present work, the application of this thesis is through a European lens. It is possible that some of these viruses may originate within Europe and merely have accumulated within these species during their stay here. Other viruses may have arrived in Europe with the plants themselves.

The import of certain ARTCs into the UK has already been studied as a potential pathway by which viruses may be introduced, specifically, ulluco (*Ullucus tuberosus*) and yacón (*Smallanthus sonchifolius*). Ulluco was the topic of an extensive commodity risk assessment by the EU, determining that ulluco is a host for Andean potato latent virus (APLV) and potato virus T, with PVT being the most likely to be found infecting these plants, at a possible 843 and out of every 10,000 tubers being infected (Bragard et al., 2021, EPPO, 2023e). In addition, a novel virus, Ullucus polerovirus 1, with serological cross-reactivity to potato leafroll virus (PLRV) and tymoviruses with serological cross-reactivity to Andean potato latent virus (APLV) were discovered in UK plantings of ulluco grown from internet purchased material (Fox et al., 2019). APLV is an A1 quarantine virus in the EPPO region (EPPO, 2023e). In another study, potato yellowing virus, also an A1 pathogen, was detected from yacón growing in the RHS botanical gardens of Hyde Hall (Silvestre et al., 2020, EPPO, 2023e). The plants were subsequently destroyed. Yacón and ulluco are currently prohibited for import in the UK (Barker, 2022, DEFRA, 2024). Ulluco is specifically prohibited due to the unknown risk posed by Ullucus tymovirus 1 and Ullucus tymovirus 2 (Ian Adams, pers comm). Exchanging germplasm across national borders may contravene the Nagoya protocol in some instances, as these crops trace their ancestry to South America, not Europe.

Table 1.1 Summary of English common names of ARTCs.

Smallanthus sonchifolius	*Polymnia sonchifolia*	Yacón	Yacon strawberry, jíquima, ground apple	*Asteraceae*	(Vitali et al., 2015; Gurung et al. 2018; National Research Council, 1989)
Ullucus tuberosus		Ulluco	Papa lisa, melloco	*Basselaceae*	(Condori et al., 2008; National Research Council, 1989)
Canna edulis	*Canna achiras*	Achira	Edible canna, purple arrowroot, Queensland arrowroot	*Cannaceae*	(National Research Council, 1989)
Arracacia zanthorrhiza		Arracacha	Peruvian parsnip, Peruvian carrot, White carrot, racacha	*Apiaceae*	(Hermann & Heller, 1997; National Research Council, 1989)
Lepidium meyenii	*Lepidium peruvianum* *	Maca	Pepper grass, pepper weed, Peruvian ginseng (marketing term)	*Brassicaceae*	(Hernándo Bermejo & León, 1994; Hermann & Bernet, 2009)
Pachyrhizus ahipa		Ahipa	Andean yam bean	*Fabaceae*	Hermann, 1997, Al-Shehbaz, 2010

Scientific name	Synonyms	Common name	Other names	Family	References
Oxalis tuberosa		Oca	New Zealand yam	*Oxalidaceae*	(Arbizu et al. 1997)
Tropaeolum tuberosum		Mashua	Isaño, Isañu, añu, mashwa	*Tropaeolaceae*	(Condori et al., 2008; Delhey et al., 1977; Johns, 1981; Hernándo Bermejo & León, 1994)
Mirabilis expansa		Mauka	Chago	*Nyctaginaceae*	(Hernándo Bermejo & León, 1994)

ARTCs, as their name suggests, originate from the Andes, sharing their centre of origin with *Solanum tuberosum*, the potato (Flores et al., 2003). The International Potato Centre (Centro Internacional de la Papa, CIP) (https://cipotato.org/roots-and-tubers/) lists nine species of ARTCs, and George, 2011 list 8 'Andean tuber crops', excluding ahipa (*Pachyrhizus spp.*) and yacón, but including bitter potato (*Solanum x juzepczukii* and *Solanum x curtibolum*). For the purposes of this thesis, nine plants will be referred to as ARTCs, summarised in Table 1.1 on Page 7. The synonym *Lepidium peruvianum* continues to be used in error, owing partly to confusion and partly to marketing; *Lepidium meyenii* remains the preferred species name used by the Instituto Peruano de Productos Naturales of Peru (Hermann and Bernet, 2009). The maintenance of considerable genetic variety within these crops and the practice of inter-cropping different species and landraces is likely a way of ensuring that, should one landrace or entire species fail to produce, at least one landrace or species planted in a field will yield at harvest due to their differing tolerance to stressors (Hernándo Bermejo and León, 1994). In particular, the plants oca (*Oxalis tuberosa*), mashua (*Tropaeolum tuberosum*), yacón and ulluco will be focused on in the present work, with the majority of the thesis dedicated to oca.

Despite their diversity, ARTCs share some features in common. The most obvious is that the primary reason to grow them is for their roots, hypocotyls, tubers and rhizomes, though their leafy material also serves as animal feed (Hermann, 1997). Most are propagated vegetatively except for maca (*Lepidium meyenii*) and ahipa (*Pachyrhizus ahipa*) which are grown from seed (Hermann, 1997). A model in cassava and sweet potato suggests that viral control practices in vegetative crops may inadvertently select for higher virus titre (and thus symptoms) in plants (Madden et al., 2000). Clonal propagation removes sexual barriers to vertical transmission when

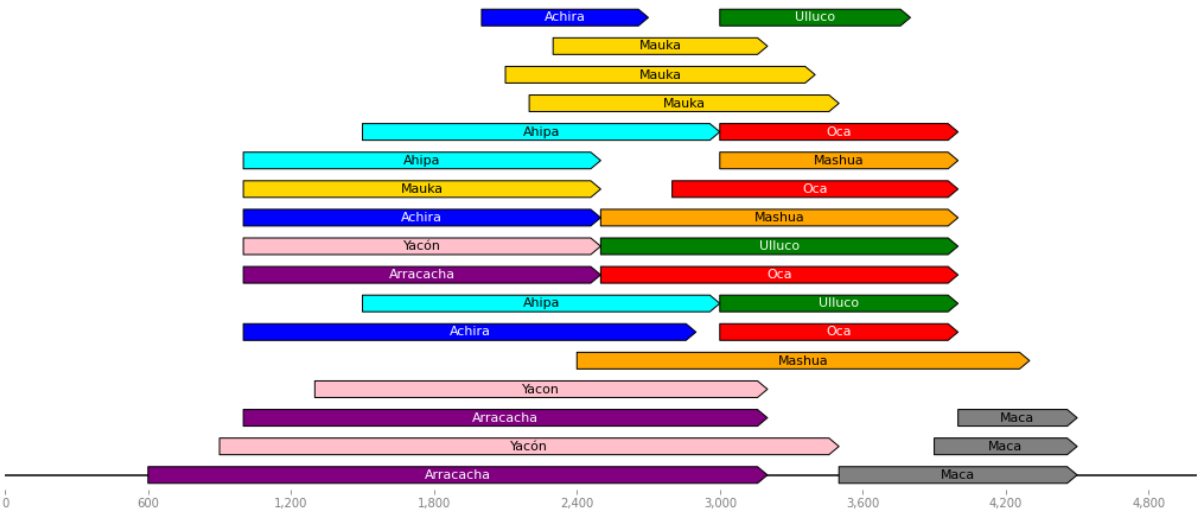


Figure 1.1 Diagram of reported elevations ARTCs are grown at in their native Andean region. Some of these crops are productive at sea level as well. Sources used are: (Flores et al., 2003; Hermann, 1997; Arbizu et al., 1997; National Research Council, 1989).

compared with seed (Sastry, 2013a). However, there are some cases where virus transmission in vegetatively propagated potato is not 100%. A study of tomato spotted wilt virus in potato plants found that no daughter plants originating from 30 infected potato tubers tested positive for the virus (Abad et al., 2005). A test of tubers from potato plants infected with tobacco rattle virus found that infected plants did not always lead to infected tubers, though the finding of an infected daughter tuber from a mother plant that tested negative maysuggest that the result was due to uneven systemic transport within the plants (Crosslin et al., 1999).

Most ARTC propagules can survive for quite some time, depending largely on the means of storage (García and Cadima, 2003). Yacón, arracacha (*Arracacia xanthorrhiza*) and mauka (*Mirabilis expansa*) cannot be propagated from the edible part of their roots (though yacón can still be grown if the storage root is still attached) (Hermann, 1997). This means that plants sold for eating should be less likely to divert to the 'plants for planting' stream. In addition, yacón roots and arracacha are quite perishable (Castro et al., 2012, Hermann, 1997). This is of consequence to their plant health risk in EPPO countries. In particular, e-commerce in plants between private individuals favours small, less perishable items, including bulbs, corms, cuttings and tubers (IPPC Secretariat, 2023a). It is worth noting here that the oca stem cuttings purchased for this thesis were not viable when planted even very shortly after arrival.

In terms of climate, the ARTCs split into three main groups: those grown in the warmest subtropical valley of the Andes (arracacha, achira, yacón, mauka and ahipa), those grown in more cool, temperate valleys alongside potato (ulluco, oca and mashua) and lastly maca, grown in the cold highlands (Puna). There is overlap between these regions (Arbizu et al., 1997, Hermann, 1997) except for maca, which in some sources shares its altitude only with some types of bitter potato (sometimes considered an ARTC, but not for the purposes of this thesis) (Hermann and Bernet, 2009). However, the summary (Fig 1.1) shows that when altitudes are taken from some sources, they produce an overlap with the upper ranges of other crops.

Different countries have different altitude limits for different ARTCs, influenced by climate (see Figure 1.2). Arracacha grows best at elevations of 1800m to 2500 in Colombia, but between 1000 and 2000 in Brazil (National Research Council, 1989), while mauka grows at different elevations depending on the country and its climate, due to lack of frost tolerance (Seminario, 2004). Of interest here is the potential 'daisy chain' formed by plants in the upper and lower extremes of their range. Depending on the country, a line of cultivation could extend from sea level to near the level of permanent ice on the Puna.
Claims about the importance of viral infection in Andean root and tuber crops are contradictory. Some claim that while virus infection is common, it has little detrimental effect to the plants (Hermann, 1997). Other studies have found that viral infection, especially mixed infections, greatly reduce the yield (Guimaraes and Flores, 2005). The National Research Council in 1989 mentioned reducing viral load (ten thousand years' worth of viral accumulation) as a necessary step in increasing cultivation of ARTCs, though mention that some take as many as ten generations to accumulate virus post-cleaning (National Research Council, 1989). Mauka shows promise as having antiviral properties, especially against PVX, PVY and potato tuber spindle viroid due to a protein it shares with *M. jalapa* (Seminario, 2004).

1.5. Cultivation Outside of the Andes

Maca Maca gained some notoriety as a health supplement in Europe, the USA and Japan in the 1990s, largely for its alleged benefits to human libido and sexual health. Peru became an exporter of dried, milled or extracted roots to supply a global market for flour and extracts at this time (Hermann and Bernet, 2009).
Maca was approved for use and cultivation in China in 2011, and is grown widely there (Yin et al., 2019). Small-scale plantings in the form of greenhouse experiments and seeds for sale have been reported in the USA, the Czech Republic and Australia (Hermann and Bernet, 2009, Valentová and Ulrichová, 2003). Maca has been the centre of discussions about bioprospecting and biopiracy, with companies outside of Latin America selling maca extracts for profit without proper consideration to the people of their countries of origin. Perú's indigenous peoples worked together to sue Pure World Inc., a US company, for exploiting and patenting maca extracts, the use of which were already known to the indigenous peoples of the Andes and are not new (Landon, 2007).

Oca Oca is also grown in Mexico, likely for more than two centuries (National Research Council, 1989). The plants have been grown in Aoteoroa-New Zealand since as early as 1869, but only gained commercial success later, likely between the 60s and 70s (National Research Council, 1989, Vietmeyer, 1999). By 2006 oca was worth $2 million per annum in the country (Martin et al., 2005). There have been experimental plantings of oca in Japan, such as a study in 1999 (Sato et al., 1999).

Ulluco and arracacha Along with oca, ulluco and arracacha were imported into Aoteoroa-New Zealand in 2005 (Fletcher et al., 2005). Plantings of ulluco in Europe (France, Finland) and Aoteoroa-New Zealand are thought to be scarce due to incompatible daylight hours and low yields, but Peruvian exports have reached the Netherlands, France, Italy, Spain, Germany, Belgium and the United Kingdom (the latter two received no exports after 2018) (Bragard et al., 2021). Ulluco has also grown successfully in Sri Lanka (National Research Council, 1989).

Yacón Yacón was brought to Aoteoroa-New Zealand from Ecuador in the 1960s, from there to Japan likely in 1985 (Koike et al., 2009), from Japan to Brazil in 1989-1991, to the Czech Republic in 1994, South Korea in 1998, Taiwan in 2000, and to China and the Philippines in 2003 (de Almeida Paula et al., 2015, Ojansivu et al., 2011). It is also grown in small to very small croppings in Russia, Mexico, Estonia, Germany, Austria and the United States (Manrique, 2005, Fernández et al., 2006, Kamp et al., 2019). Yacón was grown in Italy between 1927 and the Second World War (Ojansivu et al., 2011). Japanese companies have developed a small number of commercial varieties of yacón (Gurung et al., 2018). In Brazil one of its common names translates to “diet potato”, reflecting a global usage of the plant as part of calorie restrictive diets, or part of diet-based management of type 2 diabetes (de Almeida Paula et al., 2015, Caetano et al., 2016).

Achira Achira is grown in tropical Asia perhaps more than in the Andes, especially Vietnam (introduced by French colonists in the late 19th century and now mostly eaten as noodles(Le et al., 2013)). It is grown also in China, Taiwan, Thailand, the Philippines, St. Kitts, Indonesia, Hawai’i, Africa [1] and Burma. Lastly it is grown in Australia, primarily for starch and feedstock and in India for its rhizomes as food (Reddy and Reddy, 2015, International Potato Centre, 1999, Thitipraphunkul et al., 2003, Ugent et al., 1984). Japan has developed a ’green’ cultivar (Thitipraphunkul et al., 2003).

Ahipa Ahipa is not widely grown unlike its relative jicama, but has been successfully grown under experimental conditions in Denmark and Spain (National Research Council, 1989, Leidi et al., 2003).

Mashua Mashua has been grown as an ornamental nasturtium in Britain and the United States, and has produced good tuber yields in Aoteroroa-New Zealand, the UK, the Pacific Northwest of the USA and Canada, but they are strongly short-day limited, which may limit widespread adoption (National Research Council, 1989, Sperling and King, 1990).

Mauka Mauka plants did not flower or tuberise when grown in Mexico as an experimental planting (Seminario, 2004).

[1] Sources are not specific on which countries of what is a very large continent produce the crop, but one would presume places with a tropical climate favourable to the plant.

1.6. Online Information

Anecdotal evidence from a grower, William Whitson in the United States suggests ulluco is grown there by hobbyists in Oregon and the Appalachians, as well as elsewhere in the United States. This grower also offers virus-cleared oca and mashua tubers, yacón and mauka for sale, and has experimented with maca and ahipa seeds. The proprietor once offered arracacha and achira for sale but no longer does (https://www.cultivariable.com accessed 23/1/2024). Whitson has commented on viral findings in oca before and commits voluntarily to sharing virus-free material. The same perhaps cannot be said of the material circulating on European eBay.
A blog entry by Owen Glyn-Smith aka Radix from 2009 claims that the first truly European oca tubers were bred from seed in Wales under the name 'Pink Dragon'. One commenter claims to have found oca plants for sale in the British supermarket chain Waitrose (Glyn-Jones, 2009). An informal alliance of European breeders called the Oca Breeders Guild operated from 2015, spearheaded by Owen Glyn-Smith; by December 2022 the guild was on hiatus and no longer offered seed or tubers. The website is presently unavailable and only accessible via archive.org's 'Wayback Machine' (http://www.ocabreeders.org/).
The Open Seed Sharing Initative (https://osseeds.org/) has three varieties of mashua, twelve varieties of oca (eight of which originate from the Oca Breeders Guild) and five ulluco varieties. A number of these were pledged by William Whitson, mentioned above.

1.7. Day Length of ARTCs and Potato

As they all originate from the same region, it remains a question why potato gained worldwide significance centuries ago and the other Andean roots and tubers did not. The potato was probably first seen in Spain some time in the mid to late 1500s. Historically, potatoes do not preserve well when fresh, but the labouriously freeze-dried form was taxed alongside maize in Incan times (McNeill, 1999). Oca and ulluco freeze into *khaya* and *llingli* respectively, while all three products produce flour when ground (Bradbury and Emshwiller, 2011, Smith, 2012, National Research Council, 1989). Mashua stores well below ground (Clark, 1999).
Andean and Chilean potatoes form two distinct groups with different day-length adaptations, with their origin remaining unclear; theories include a hybrid origin as potato cultivation moved down the Andean slopes to Southern Chile, or separate origins for the two groups (Gavrilenko et al., 2023, Hardigan et al., 2017). Early European potatoes were chiefly Andean in origin and later, day-length adapted potatoes from the late 18th century until the blight are likely descendants of either a) *de novo* mutants that emerged due to extensive European breeding and hybridisation or b) plausibly derived via introgression from daylight-adapted Chilean potatoes that became more common in Europe after the late 18th century (Gutaker et al., 2019).
Exchange of germplasm from the Americas *to* Europe in the 16th century seems to have taken place in a somewhat chaotic and disorganised manner through private enterprise, the Iberian authorities being chiefly concerned with the extraction gold and other valuable metals (Hernándo Bermejo and León, 1994). Potatoes gained popularity in Europe only later, from the

mid 18th century, in part due to alterations in agriculture that allowed potatoes and maize to enter the traditional crop rotation; the potato is credited with enabling explosive population growth in Northern Europe between 1750 and 1850 (De Jong, 2016).
Ahipa, arracacha, achira and yacón are all either day length neutral or adaptable, but none are widely grown in the Andes (Hermann, 1997, Laitenberger, 2018). Arracacha, while nourishing, is highly perishable (Hermann, 1997). Achira and yacón may have a worldwide spread at present, but are not efficient enough food sources to be staples as potato now is (Gade, 1966, Ulloa et al., 2021). Maca requires a long fallow period to be productive and it, like mauka and ahipa, was a relative unknown outside of specific areas until relatively recently (Hermann and Bernet, 2009, Gendall et al., 2019, Hermann, 1997).
The nutritionally dense, easily stored, highly productive tubers of oca, mashua and ulluco compare to potato in many respects. The three are widely cultivated alongside potato in the Andes as staples and have been since ancient times (National Research Council, 1989). However, these other comparable Andean root and tuber crops do not have such the secondary centre of diversity outside of the equatorial highlands that potato does (García-Díaz et al., 2023). Oca cultivation in other countries has largely been concerned with breeding cultivars adapted to tuberise in a broader day-length window, thus avoiding freezing of the foliage during Winter before a suitable harvest was obtained (see section 1.6). As mentioned above, ulluco saw little success in the recent past due to this issue (Bragard et al., 2021) and this factor was proposed to be a limit to widespread adoption in mashua (National Research Council, 1989). Possibly the pungent mashua is simply not as palatable as the others, and may be abandoned when other food sources are available (Clark, 1999). Some oca lines were readily accepted in taste tests, however (Sangketkit et al., 2000).
The reason for this day-length disparity between potato and the other staple Andean tuber crops hinges in part on the true origin of the Andigena and Chilotanum groups and associated wild ancestors in the Andes and Southern Chile, which remains to be conclusively determined. Possibly, the lack of a wild ancestor of ARTCs in the Southern parts of Chile might have prevented a day-length adapted regional variant becoming available to early modern Europeans.

1.8. Details of Oca

Oca is the main host studied in this thesis. It has been documented in Europe since at least the 19th century (Chirat, 1841). In order to understand the life history of its viruses (and thus the risk they may pose) it is important to understand the life history of oca. Unfortunately, its genome is not well characterised; as will be seen, this can make studying DNA viruses and their associated endogenous viral elements (EVEs) a challenge. On the positive side, it is a popular crop in its native Andean highlands, second only to potato in production and consumption (National Research Council, 1989). For this reason, it has been studied in some detail for its importance to those who grow it, and its potential in industry.
Oca is a member of the family *Oxalidaceae*, sometimes known as the wood sorrels or false shamrocks. Members of its genus grow both as wild plants native to Britain (*O. acetosella*,

common wood sorrel, being the only native British species), and naturalised introductions such as *O. corniculata* and *O. corymbosa* commonly found growing in disturbed habitats (Young, 1958). A database of plants in Europe lists thirteen 'naturalised aliens'. *O. decaphylla* and *O. fontana* are listed as 'casual aliens' or 'aliens (unknown)' in the British Isles (Henning and Raab-Straube, 2016). There are over eighty species of *Oxalis* in Peru alone, 39 of which are endemic (Brako and Zarucchi, 1993).

As a domesticated crop, oca originates from the valleys of the Andes, between 2,500 and 3,800 masl, alongside ulluco and mashua. Oca plants tuberise with an optimum daylight time of 11-12 hours (King, 1987), though Aoteoroa-New Zealand, Mexican and possibly Chilean lines show promise for being day-length neutral (National Research Council, 1989). Varieties were developed in Aoteoroa-New Zealand (a temperate climate) that produce marketable tubers in June (Winter), though the tubers have been grown commercially there for well over a century (Martin et al., 2005). Oca is grown through vegetative propagation, though introduction of seed-grown plants has probably shaped the species' diversity in the field (King, 1987, Bonnave et al., 2016). Oca tolerates poor and marginal soils well, requiring little care to produce a crop. Stolon formation begins at four months after planting, where the traditional method of hilling begins, much like potato. Following two months of this, the tubers mature for 2-3 before harvesting, though some Mexican varieties are harvest-ready in six months from planting (National Research Council, 1989).

It is unknown whether the tubers circulating in the European online market stem most recently from Australasia or the Americas. Genetic diversity studies found the South American plants to split roughly into Peruvian and non-Peruvian (other South American) lines, with low genetic variation attributed to their clonal propagation method (Pissard et al., 2006). Which of these lines are closer to the European lines is not yet known.

In their native Andes oca is inter-cropped with potato to aid in pest management practices (Sperling and King, 1990). Oca is grown widely in the Altiplano region, and work is ongoing to increase yields to provide income for the farmers there, many of whom live in poverty (Leon Velarde et al., 2008).

In Bolivia oca is grown as food and to sell both 'seed' (tubers) and edible tubers at market (Bonnave et al., 2016). The 2016 study by Bonnave et al., composed of interviews with farmers, found that most farmers faced yield reduction year on year due to 'tiring' of seeds and would need to buy replacements from another region to compensate - but not all. This is similar to potato, where the process of successive generations reducing in quality is called 'degeneration'. In the UK potato growing sector, seed potatoes are certified based on their high quality (free of disease) and are then bulked up for culinary or industrial use. The majority of the 'yield gap' (difference between observed and expected yields) in potato is due to disease infestation, which the seed certification sector spends considerable time and finance to prevent (Forbes et al., 2020).

1.9. Viruses with Andean Root and Tuber Crop Hosts

Table 1.2 A table of the virus species which were first described infecting an ARTC host, including the virus abbreviation, whether at least one complete representative sequence is available (and which year) and whether the virus was first reported via HTS. Under the Seq column, representing sequence information, a virus abbreviation followed by a question mark indicates where a virus may have been sequenced under a new name.

Arracacha latent virus C	ALVC	*Crinivirus*	Arracacha, 2021 2021	Y	De Souza et al. 2021
Arracacha latent virus E	ALVE	*Enamovirus*	Arracacha, 2021 2021	Y	De Souza et al. 2021
Arracacha potyvirus 1	AP-1	*Potyvirus*	Arracacha, No 1997	N	Hermann, 1997
Arracacha virus 1	ArrV-1	*Closterovirus*	Arracacha, 2018 2018	Y	Orílio et al. 2018
Arracacha virus 2	AV2	*Unknown*	Arracacha, No 1994	N	Santa Cruz and Jayasinghe, 1994
Arracacha virus 3	AV-3	*Carlavirus*	Arracacha, No 1997	N	Hermann, 1997
Nepovirus arracaciae	AVA	*Nepovirus*	Arracacha, 2017 1978	N	Jones, 1978; Adams et al. 2017
Cheravirus arracaciae	AVB	*Cheravirus*	Arracacha, 1979	N	Kenten and Jones, 1979
Arracacha virus V	AVV	*Vitivirus*	Arracacha, 2017 2017	Y	Oliveira et al. 2017
Arracacha mottle virus / Arracacha virus Y	AMoV / AVY	*Potyvirus*	Arracacha, 2013 (as 1990 AMoV) ***	Y	Brunt, 1990; Orílio 2013; Adams et al. 2018
Papaya mosaic virus	PapMV	*Potexvirus*	Ulluco, 2019 1982	N	Brunt et al. 1982; Fletcher, 2001; Fox et al. 2019
Nepovirus solani	PBRSV	*Nepovirus*	Arracacha, No 1994	N	Lizárraga et al. 1994
Mashua virus Y	MVY	*Potyvirus*	Mashua, 2018 2018**	Y	Adams et al. 2018
Mirabilis potyvirus 1	Mir-1	*Potyvirus*	Mauka, No 2000	N	Lizárraga, 2000
Tropaeolum mosaic virus	TropMV	*Potyvirus*	Mashua, MVY? 1998	N	Soria et al. 1998
Tropaeolum tuberosum potyvirus 1	Not given	*Potyvirus*	Mashua, MVY? 1990	N	Brunt et al. 1990
Tropaeolum tuberosum potyvirus 2	Not given	*Potyvirus*	Mashua, MVY? 1990	N	Brunt et al. 1990

Species	Abbr.	*Genus*	Host	Seq	HTS?	Ref.
Ullucus comovirus 1	UCoV1	*Comovirus*	Ulluco, 2019	2019	Y	Fox et al. 2019
Ullucus mild mottle virus	UMMV	*Tobamovirus*	Ulluco, 1982	UTobV1?	N	Brunt et al. 1982; Fletcher, 2001; Offei et al. 1995
Ullucus mosaic virus	UMV	*Potyvirus*	Ulluco, 1982	UPotyV1?	N	Brunt et al. 1982; Fox et al. 2019
Ullucus polerovirus 1	UPolV1	*Polerovirus*	Ulluco, 2019	2019	Y	Fox et al. 2019
Ullucus potyvirus 1	UPotyV1	*Potyvirus*	Ulluco, 2019	2019	Y	Fox et al. 2019
Ullucus tobamovirus 1	UTobV1	*Tobamovirus*	Ulluco, 2019	2019 UMMV?	Y	Fox et al. 2019
Ullucus tymovirus 1	UTyV1	*Tymovirus*	Ulluco, 2019	2019	Y	Fox et al. 2019
Ullucus tymovirus 2	UTyV2	*Tymovirus*	Ulluco, 2019	2019	Y	Fox et al. 2019
Comovirus ulluci	UVC	*Comovirus*	Ulluco, 1982	UCoV1?	N	Brunt et al. 1982; Fox et al. 2019
None	N/A	*Cytorhabdovirus*	Arracacha, 2016	2016	Y	Gomes et al. 2016
Badnavirus maculas-mallanthi	YNMoV	*Badnavirus*	Yacón, 2015	2015	Y	Lee et al. 2015
Yacon virus A	Not given	*Capillovirus*	Yacón, 2015	2015	Y	txid1868472
Yam bean mosaic virus	YBMV	*Potyvirus*	Ahipa, 2012	2012	Y	Fuentes et al. 2012

Table 1.3 A table of virus species whose members have been detected in ARTCs, but were not initially isolated from them. This thus excludes examples where the virus found infecting the plant was a new strain, such as AVB-O and PapMV-U.

Oca	*Nepovirus solani*	Jones and Kenten, 1981	Peru	Experimental infection by Lizárraga et al, 2000
	Potato virus T	Lizárraga et al. 2000		
	Papaya mosaic virus	Fletcher and Fletcher, 2001	Argentina	
	Ullucus mild mottle virus	Fletcher and Fletcher, 2001	Bolivia	
	Nepovirus arabis	De Jonghe et al., 2022	Europe	
Oca	*Ullucus polerovirus 1*	De Jonghe et al., 2022	Europe	

Host	*Species*	Paper	Origin	Notes
Yacón	*Potato yellowing virus*	Silvestre et al. 2020	Peru, Ecuador, UK, Bolivia	
	Cucumber mosaic virus	De Jonghe et al., 2022	Europe	
	Pepino mosaic virus	De Jonghe et al., 2022	Europe	
	Caulimovirus tessellodahliae	De Jonghe et al., 2022	Europe	
Maca	*Tymovirus latandigenum*	Kreuze et al., 2020	South America	
	Turnip mosaic virus	Yin et al., 2015	China	Experimental infection by Marthe et al., 2003
	Tymovirus mosandigenum	Kreuze et al., 2020	South America	
Ulluco	*Potato virus T*	Lizárraga et al., 1999	South America	
	Tymovirus latandigenum	Bragard et al., 2021	South America	Other detections may not be APLV
	Tymovirus mosandigenum	Kreuze et al., 2020	South America	
	Potato leafroll virus	Bragard et al., 2021	South America	Other detections may not be PLRV
	Nepovirus arracaciae	Lizárraga et al., 1999	South America	
	Fabavirus alphaviciae	Fox et al. 2019	UK	
	Fabavirus betaviciae	Fox et al. 2019	UK	
	Orthotospovirus tomatomaculae	De Jonghe et al., 2022	Europe	
	Turnip yellows virus	De Jonghe et al., 2022	Europe	
	Potato virus S	Kreuze et al., 2020	South America	
Arracacha	*Potato virus S*	Brunt, 1990*, De Souza et al., 2018	Peru	
	Nepovirus solani	Lizárraga et al., 1994	South America	
	Bidens mosaic virus	Orílio et al., 2017	Brazil	
Achira	*Cucumber mosaic virus*	Lizárraga, 2000	Peru	
	Banana streak virus (species unknown)	Lizárraga, 2000, Reichel, 1997	Peru, Colombia	
	Tomato aspermy virus	Lizárraga, 2000	South America	
	Bean yellow mosaic potyvirus	Lizárraga, 2000	South America	
	Badnavirus alphamaculaflavicannae	Lizárraga, 2000	South America	
	Sugarcane mosaic virus	Betancourth, 2020	Colombia	
Mauka	*Caulimovirus tessellomirabilis*	Lizárraga, 2000	South America	
Mashua	*Potato virus T*	Lizárraga et al. 2000	South America	

Host	*Species*	Paper	Origin	Notes
	Potato leafroll virus	National Research Council, 1989		
	Sowbane mosaic virus	Reers, 2016	Unknown	
	Verbena latent virus	De Jonghe et al., 2022	Europe	
	Red clover (vein) mosaic carlavirus	De Jonghe et al., 2022	Europe	
	Alphanucleorhabdovirus physostegiae	De Jonghe et al., 2022	Europe	
	Turnip yellows virus	De Jonghe et al., 2022	Europe	
Ahipa	*Bean common mosaic virus*	Sørensen, 1996	South America	

ARTCs are known hosts of a number of viruses, some of which are economically important. Table 1.2 on Page 15 summarises the viruses first discovered on these plants and reported in the literature. Other, known viruses, have also been found infecting them, which are summarised in Table 1.3 on Page 16.

The likely first new virus report in ARTCs comes from 1977 (Delhey et al., 1977). While studies in the 1980s suggested that the indigenous viruses present in ulluco could not infect potato (Brunt et al., 1982b), ulluco plants infected with APLV were later found in mixed plantings with potato (Lizárraga et al., 1996b). Lizárraga et al. suggest that, not only are known potato viruses found within ARTCs, but that viruses within these crops could evolve to infect potato, and vice versa. Three main viruses within ulluco can already infect potato: APLV, PLRV and potato virus T (PVT) (Lizárraga et al., 2001). PVT has also been found to infect oca and mashua naturally, with minimal symptoms (Lizárraga et al., 2000). The virus is not currently known to be present in the UK and is an A1 regulated pest. Pertinent to this thesis is the existence of the same virus infecting multiple ARTCs despite their botanical diversity.

Table 1.3 on Page 16 summarises reports of viruses from ARTCs that are *not* the first isolation host. It *does* include examples where a virus strain not originally isolated from one ARTC was found on another (e.g. arracacha virus A in ulluco and papaya mosaic virus ulluco strain (PapMV-U) in oca). This differs from the previous table by excluding first findings, thus only including viruses which have at least two known hosts, including in some cases more than one ARTC. That is, if a virus was reported in arracacha only (i.e. the isolation host), it is on the first table but not the second.

Crops grown at the same elevation share some viruses. PVT is common to both wild and cultivated potato, oca, ulluco and mashua, for example (Salazar and Harrison, 1978). PVT has no known vector, and nor does the genus as a whole (Russo et al., 2009, Xylogianni et al., 2021); vegetative propagative material is listed as the main cause of spread in the papers cited here, but clearly there must be some means of horizontal transmission or there would be no natural hosts within the ARTC group. Ulluco, mashua and potato are all reported hosts of PLRV, though as has been noted, it is possible that some of these reports actually reflect a finding of a serologically related species.

Blueberry leaf mottle virus has been reported in oca, but studying the data revealed it to be a novel nepovirus. Similarly, reports of Fragaria chiloensis latent virus are likely to be potato yellowing virus, common in yacón (see section 1.4), so they have been excluded from Table 1.3 (De Jonghe et al., 2022).

In a study in ulluco plants from the UK, some viruses were detected by ELISA before HTS revealed another, closely related virus within the bulk (Fox et al., 2019). For example, plants forming the Ullucus polerovirus 1 bulk reacted to PLRV antisera. Though this virus is likely its own species based on genomic information, it is possible that this virus represents the 'PLRV' isolate from ulluco reported in 1996 (Lizárraga et al., 1996a). Similarly, plants from the bulk containing two novel tymoviruses cross-reacted with APLV antisera but are likely to be different species based on genomic information (Fox et al., 2019). APLV was reported in ulluco plantings based on serological information in South America before (Lizárraga et al., 1996b).

A later risk assessment of ulluco reported that lab-based methods conducted by SENASA, including PCR, detected PLRV and APLV in ulluco, but the paper did not clarify if PCR or ELISA was used to detect APLV and PLRV (Bragard et al., 2021).

Ullucus tobamovirus 1 may be synonymous with what was initially reported as *Tobacco mosaic virus*, ulluco strain, later to be known as *Ullucus mild mottle virus*. *Ullucus comovirus 1* may be synonymous with *Comovirus ulluci* (Ullucus virus C), while *Ullucus potyvirus 1* may be synonymous with *Ullucus mosaic virus*; sequences were not reported for these species when they were first described (Fox et al., 2019).

Mashua virus Y could be synonymous with any one of *Tropaeolum mosaic potyvirus*, *Tropaeolum potyvirus 1* or *Tropaeolum potyvirus 2* and this was noted in the paper announcing its discovery (Adams et al., 2018c). As noted before, Lizárraga et al. suggested that the long-term association of these species with one another has facilitated transfer of their viruses. If the findings of Ullucus tymovirus -1 and -2 ,and UlPolV1 are most similar to the isolates of APLV and PLRV from the two 1996 papers, this might imply that the hypothesised jump between host species happened long enough ago for speciation to occur in the interim.

Sequence named as a member of a new species, *Yacon virus A* was submitted to GenBank as accession KU375548.1. A phylogenetic study of this virus and other capilloviruses found that yacon virus A, along with the other members of its genus, originates from outside of South America; tests on yacón plants from South America did not find the virus (Shujaei et al., 2020).

Arracacha latent carlavirus was also reported but is a synonym of *Potato virus S* (Santillan et al., 2018a). *Arracacha latent carlavirus* has previously been mentioned as a possible synonym of *Arracacha virus 3*. Possibly, *Arracacha potyvirus 1* and *Arracacha virus Y* are synonymous as well, which could mean three names for this species (alongside *Arracacha mottle virus*) (De Souza et al., 2021, Adams et al., 2018b). *Arracacha virus 2* was described in 1994 according to a secondary source, but the original paper was not available to this author and the genus is thus not known (Henz, 2002, Santa Cruz and Jayasinghe, 1994).

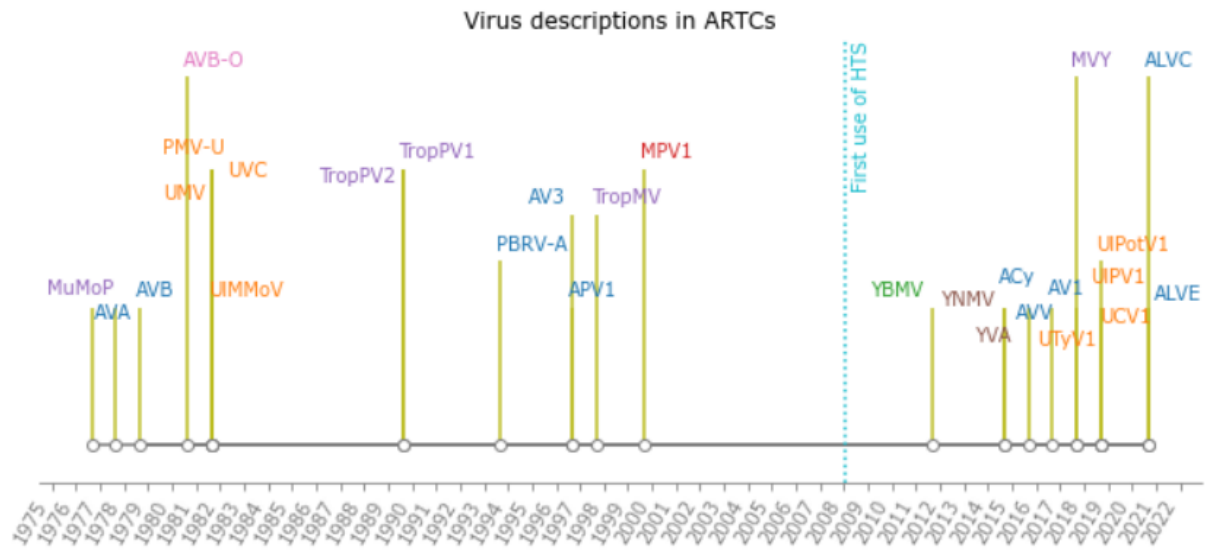


Figure 1.2 A timeline of novel virus findings in all ARTCs coloured according to the isolation host. Purple = mashua, green = ahipa, blue = arracacha, orange = ulluco, red = mauka, brown = yacón, pink = oca. The first use of HTS in plant virology is added to indicate the pre- and post- HTS discoveries.

1.10. Timeline of Novel Virus Discoveries in Andean Root and Tuber Crops

Two viruses are discussed here that were not included in the table above as no name was ever proposed for them to be ratified. In 1977, virus particles likely to be *Potyvirus* members from plants with mosaic symptoms were reported in mashua but not named; this could be any of the named *Potyvirus* species or a new species (Delhey et al., 1977). The unnamed tentative cytorhabdovirus (initially detected by HTS) was also included from a study in 2016 (Gomes et al., 2016).

When accounting for the fact that *Arracacha mottle virus* (Orílio et al., 2009) is synonymous with *Arracacha virus Y* (Adams et al., 2018b), there are two large gaps in the timeline concerning novel virus discovery in these plants. The first gap is smaller, a gap of eight years between 1982 and 1990. The second is a slightly longer gap of twelve years between 2000 and 2012. The studies can be brigaded into three main groups: two broad efforts in the 20th century (and early 21st century), and the current ongoing body of work, which is less centralised.

1.10.1. 20th Century & Early 21st Century

The Americas As these crops originate in South America, general studies of the crop are excluded for brevity. The information most relevant to virology is summarised here. The first tentative novel virus (an unnamed member of the genus*Potyvirus*) was reported in mashua in 1977, likely to be the first finding in any ARTC; the sample was collected from a diseased plant in Bolivia in 1964 (Delhey et al., 1977). Much virological work was conducted in the late 1970s and early 1980s in South America proper. Roger Jones, working at the CIP, collected arracacha from Huanuco in Peru (Jones and Kenten, 1978, Kenten and Jones, 1979), oca from Huasahuasi in Peru (1975 sampling) (Jones and Kenten, 1981), and ulluco from Bolivia and Peru (Brunt et al., 1982b, Brunt et al., 1982a). In 1995, Brunt, one of the original authors of the *Ullucus mild*

mottle virus paper, along with others, did confirmatory work on what was originally called tobacco mosaic virus, ulluco strain (TMV-U) (Offei et al., 1995).
In the 1980s and 1990s ARTCs became of interest to international organisations and governments for conservation and exploitation, such as the International Board for Plant Genetic Resources (Castillo et al., 1988). A 1984 seminar by the National Research Council of the United States contacted 200 plant scientists and compiled the resulting information on under-exploited Andean crops (National Research Council, 1989). This publication focused on improvement, not just for local interest, but for international and industrial use.
In the 1980s Ecuador established a department of plant genetic resources, including expeditions to collect germplasm (Muñoz and Castillo, 1991, Castillo, 1995). The Andean Network on Plant Genetic Resources (Red Andina de Recursos Fitogenéticos) emerged in the early 1990s to co-ordinate germplasm collection and improvement efforts across the Andean countries (Castillo, 1995). In addition, some studies from the USA investigated ARTCs and their viruses in the 1990s as part of an international development program (Soria et al., 1996, Soria et al., 1998).
In 1992, the CIP (part of the international organisation CGIAR) had 468 accessions of 'other andean root and tuber' crops (CGIAR Technical Advisory Committee, 1992). By 2009, this number was 2345, including uncleaned accessions, the majority of which were oca and ulluco (Panta et al., 2009). Between 1993 and 2003, the Andean Root and Tuber Biodiversity Project (Proyecto de Biodiversidad de Raíces y Tubérculos Andinos) brought together over twenty researchers under the auspices of CIP and other organisations (including the Swiss Agency for Development and Cooperation), producing a number of publications on the biology and viruses of these crops (García and Cadima, 2003). Scientists from the CIP, such as Charlotte Lizárraga, investigated the viruses in these plants, with aim to improve the propagation material available to farmers (Lizárraga et al., 1999, Lizárraga et al., 2000, Lizárraga et al., 2001, Lizárraga et al., 1996b).
Broadly, this era follows the 'Seek and ye shall find' driver of virus discovery outlined by Fox and Mumford (Fox and Mumford, 2017): commodity or institute research, often towards a broader goal. In this case, the stated goal was improvement of genetic resources.

Aoteoroa-New Zealand Oca has been grown in Aoteroa-New Zealand for upwards of 150 years (Vietmeyer, 1999). New germplasm was imported in 1993 ((Grau and Halloy, 1994) as cited in Fletcher et al., 2005 and Martin et al., 1999) and subsequently, new lines were developed from these and existing Aoteoroa-New Zealand accessions based on consumer preference tests (Martin et al., 2005). This ongoing research included imported accessions of ulluco and arracacha (Fletcher et al., 2005). Virus elimination work was performed prior to release of the accessions (Fletcher, 2001), but no novel viruses were reported during this time. Notably, the adaptation of some oca lines to the day-length of Aoteoroa-New Zealand was suggested as a plausible stepping stone to cultivation in Europe and North America, along with South Andean and Mexican lines (National Research Council, 1989).

1.10.2. 21st Century

The most recent, and ongoing, spate of research has been global, including Asia and European as well as South American and Australasian studies. The first use of HTS to find a new virus in ARTCs was in Peruvian ahipa in a field trial in 2012; with other *Pachyrhizus* species, siRNA sequencing was used to recover the genome of *Yam bean mosaic virus*, a member of the genus *Potyvirus* (Fuentes et al., 2012). In 2013, arracacha virus B, oca strain (AVB-O) infected oca material from DSMZ was sequenced in the UK through total RNA and determined to be a member of the genus *Cheravirus* (Adams et al., 2013). This is one of several studies that sequenced historical samples from the original era mentioned above (Adams et al., 2018c, Adams et al., 2018b).
In 2015, *Yacon necrotic mottle virus* was described from South Korean yacón (Lee et al., 2015). In 2016, PCR and RACE were used to recover the complete genome of an unnamed cytorhabdovirus in arracacha initially discovered via HTS in Brazil (Gomes et al., 2016) and yacon virus A, a capillovirus, was entered as GenBank accession KU375548.1. Interestingly, this accession's title reports finding Fragaria chiloensis latent virus in these samples, which is likely to be PYV. This is because PYV was not sequenced in full until 2019 (Knierim et al., 2019) thus causing incorrect assignment of PYV sequence to Fragaria chiloensis latent virus using some pipelines; the two viruses are likely to be synonymous (Silvestre et al., 2020). The data appear to originate from a Polish NPPO.
Studies in South America uncovered other novel viruses (Orílio et al., 2018, De Souza et al., 2021, Oliveira et al., 2017). Samples from an ulluco planting in the UK revealed several more, some of which may be sequence of historically isolated viruses, including possibly the 'APLV' and 'PLRV' from ulluco in 1996 (Fox et al., 2019, Jones et al., 2021). Then there is this thesis, which uses some data and preserved European samples from that time, as well as European samples from 2021-2023. This work adds ten or so viruses to the novel findings in ARTCs - the exact number requires more investigation (in particular, see **??**). This later era constitutes a wider mix of driving factors than did the previous century: agricultural improvement goals, trade concerns (especially with the UK interceptions) and, of course, application of new diagnostic materials to previous knowns and unknowns (Fox and Mumford, 2017).

1.11. High Throughput Sequencing

High throughput sequencing (HTS) was first used in plant virology in 2009, with three major papers in that year: the first used reverse-transcribed cDNA fragments to discover a novel marafivirus in grapevine (Al Rwahnih et al., 2009), the second discovered a badnavirus and a mastrevirus in sweet potato by sequencing of small RNAs produced by the host plants' defence system (Kreuze et al., 2009). The last discovered gayfeather mild mottle virus by total RNA sequencing of a virus from *Liatris spicata* inoculated onto *Gomphrena globosa* and was published shortly after (Adams et al., 2009).

High throughput sequencing in its various formats (libraries, enrichment methods and machines) has been the source of countless virus discoveries since it first emerged into plant virology in the late 2000s (Villamor et al., 2019). It is target-agnostic, requiring no expert knowledge of predicted pathogens in order to choose an assay and make a diagnosis (in theory) (Adams et al., 2009).
There remains no single method for HTS. Each method (or sub-process within a method) exists to exploit different strengths, and overcome different weaknesses, of high-throughput sequencing as it exists today. The chief split when it comes to plant virology comes from the need for enrichment of target sequence to enhance detection and reduce overheads. Viruses often exist within the cells they infect at very low quantities. Detecting low titre viruses in samples becomes difficult if host sequence is present in large amounts due to noise and dilution (Kreuze et al., 2009).

siRNA The first method utilises short interfering RNA (siRNA). When a plant becomes infected with a virus, its silencing machinery produces small RNAs of 21, 22 and 24 nucleotides in length (Massart et al., 2019). SiRNA sequencing utilises this fraction of total RNA for sequencing, for example by purification from a acrylamide gel, in order to reconstruct viral genomes of both RNA and DNA types, from different tissues within the plant (Kreuze et al., 2009). Because this method relies on the plant's immune response, sequencing of nucleic acids extracted using the siRNA method can have difficulty detecting viroids and circular DNA viruses (Vivek et al., 2020).

dsRNA Double stranded RNA (dsRNA) is another approach used, aiming to detect replicating virus complexes (Adams and Fox, 2016), though it may overlook negative sense single-stranded RNA viruses as they do not accumulate high concentrations of dsRNA during replication (Roossinck et al., 2015).

VANA Another method utilises virion-associated nucleic acids (VANA), by purifying virions from sap and then extracting the nucleic acids from them for analysis (Filloux et al., 2015a). Despite this, purification or semi-purification using VANA is able to recover non-encapsidated viruses such as endornaviruses (Ma et al., 2019).

Host sequence depletion Host depletion removes host nucleotides from a total extract (rather than selectively extract desired sequence from host sequence as above), such as total RNA (totRNA). Typically this takes the form of plant ribosomal RNA depletion via subtractive hybridisation, as rRNA forms a majority of the RNA within a cell (upwards of 90 percent in some species) (Gaafar et al., 2021, O'Neil et al., 2013). Sequencing of extracts obtained by the totRNA method may not detect dsRNA viruses as the reverse-transcription step may not reliably copy dsRNA viruses into cDNA (Roossinck et al., 2015). Removing host-derived reads via mapping as part of the after sequencing but pre-assembly can also be chosen as an *in silico*

method but relies on a well-characterised host and does not alter coverage depth in the sequencing machine (Daly et al., 2015).

1.12. The Biological Desert

The speed of discovery allowed by HTS has created a bottleneck in follow-up studies, a phenomenon dubbed the 'biological desert' during a meta-study of viruses first detected by HTS in fruit trees and their associated biological data (Hou et al., 2020). By investigating novel virus reports, this paper found that the percentage of accompanying findings drop off considerably with the effort (time, resources) required. For example, while 94% of these papers included primers for detection (a relatively inexpensive and quick endeavour), only 8% conducted host range studies (which need glasshouse space, expertise, and time). This has led to an unfortunate drought of contextual information accompanying viral discovery. In fact, a later study on 45 tomato-infecting viruses discovered in recent years noted that those which were discovered by investigating the etiology of a disease (the 'traditional' mode) were better characterised than those discovered incidentally through HTS (Rivarez et al., 2021). Many of these latter viruses received no further study after the first publication reporting their discovery. Even long before HTS, the lack of biological information on novel viruses was being lamented in 1981: "The literature contains too many examples of known viruses re-described under new names, or of viruses so inadequately described that it is doubtful whether they really are 'new'." (Hamilton et al., 1981). This paper by Hamilton et al. suggests that the following basic properties are sufficient in the 'diagnostic phase' - the particle morphology (insofar as it gives insight into taxonomic placement) and serological tests for identification. In today's virological era, taxonomic placement can be suggested from sequence, and identification can be achieved via nucleotide derived from said sequence (albeit with caveats discussed later). In some cases, such as environmental samples, the bare minimum can be achieved with barely any biological information at all.

Many plant viruses are now found through metagenomic studies with no direct association with disease (Roossinck et al., 2015). Previously, the paradigm of investigating diseases of economic importance led to a disproportionate number of viruses known to science originating from crop species (Roossinck, 2019). There are exceptions: for example, Arabis mosaic virus was discovered in *Arabis hirsuta* in an insect-proof greenhouse, having "unaccountably appeared" there (Smith and Markham, 1944). The non-targeted nature of HTS, which is theoretically (but not always in practice (Gaafar and Ziebell, 2020, Pecman et al., 2017)) unbiased, allows the study of plant viromes quickly and relatively cheaply (Maree et al., 2018). It has also allowed the discovery of viruses which may be latent, asymptomatic, or cryptic, where previously they may have gone unnoticed as their existence within plants is not readily apparent and may even be beneficial to the host (Roossinck, 2011).

As part of characterising a new virus, an update to the 2017 framework suggests some measures which might be taken when biological information is not yet available (Massart et al., 2017, Fontdevila Pareta et al., 2023). The first is a literature review of closely related species,

particularly within the same genus. From this, the framework asserts that assumptions can be made about a species in order to test them: modes of vertical (e.g. seed) or horizontal (e.g. vectors) transmission, host range (particularly how broad it is likely to be), pathogenicity and symptoms, satellites, and sources of resistance or tolerance. These must then be tested, but the assumptions allow tests to be more targeted and efficient. An oft-used method of inferring causation in plant virology aims to fulfill Koch's postulates, which has been used to associate organisms with disease since the 1880s, with a medley of revisions over the years (Evans, 1976). Bos updated the postulates for plant virology in 1981, with a focus on the ability to transmit a virus to a healthy host, for it to be multiplied therein, and for it to be purified from said host as evidence of transmission (Bos, 1981). Viruses have a number of biological qualities which differentiate them from the cellular organisms for which the postulates were first proposed. Abiotic factors, cultivar of the host, and importantly mixed infections including syngeristic ones all make causal inference a complex matter for which fulfillment of 'Koch's postulates' in a strict sense may not even be appropriate (Fox, 2020).

Some studies use full-length infectious clones as a means to circumvent the particular issue of low titre or mixed infections in plants and thereby demonstrate Koch's postulates. This requires the full genome, however, is time consuming, and lacks standardisation (Youssef et al., 2011). There are alternatives to Koch's postulates which utilise populations rather than individuals to demonstrate causation, bypassing the need for isolation entirely. For example, a study of carrot internal necrosis was able to associate symptom development with carrot yellow leaf virus using statistical methods based on an HTS survey (Adams et al., 2014).

In any case, attaching a virus as an organism with disease in an empirical fashion is central to assessing the risk an organism may pose (IPPC Secretariat, 2013). Whether that involves a complex of other species or combination of abiotic and biotic factors remains part of the assessment of said risk.

A person may consider this search for information as a form of academic triage, whereby resources are dedicated to those viruses which appear to be the most destructive. However, the phenomenon of emerging viruses would suggest that it is always valuable to understand a virus as much as possible. Emerging viruses which change host are often highly damaging (Jones, 2009). Nearly half of all emerging diseases are viral, with climate change likely to accelerate this process (Anderson et al., 2004). Though the rapidly accumulating number of viruses without biological information presents a challenge to regulators (Macdiarmid et al., 2013, Massart et al., 2017), it has allowed us to understand viral diversity in ways not possible before (Roossinck et al., 2015).

In this thesis, the risk of Andean root and tuber crops traded via the internet to the UK was investigated using high throughput sequencing, data mining of the short read archive, and a literature review of nepoviruses, as two novel members of the genus were detected during the present study. The advantages of high throughput sequencing and possible challenges to its future implementation are also interrogated more broadly, in a risk assessment context.

chapter 2 introduced the methods used for all following chapters. chapter 3 discusses viruses that were detected when oca tubers purchased from online vendors were investigated using high throughput sequencing, including a number of novel viruses from differt genera whose genome organisation is tentatively proposed. chapter 4 discusses viruses, including novel viruses detected in oca, yacón, mashua and ulluco sourced from a nursery in Ireland who expressed interest in growing these crops in a biosecure manner. chapter 5 discusses attempts to mechanically inoculate a novel nepovirus onto herbaceous hosts and contains a literature review of the host range of all known nepoviruses used to aid in constructing a panel of plants to inoculate. chapter 6 discusses the viruses detected through data mining of short read archive data derived from RNASeq of known ARTCs, in particular maca. Finally, these findings are brought together in a discussion of the threat ARTCs may pose to UK plant health, organised according to the considerations of the formal risk assessment process.

Chapter 2. Methods and Materials

2.1. Samples

2.1.1. Internet Samples for High Throughput Sequencing

Oxalis tuberosa tubers were purchased online between 2019 and 2021 from vendors on eBay.co.uk, or direct from a commercial supplier trading online in the United Kingdom (UK). Online purchases were conducted opportunistically, with the aim to investigate the material which might be purchased by a typical amateur grower in the UK.
High throughput sequencing datasets generated from a study of internet-traded tubers purchased in 2019 were also investigated. The samples whose sequences were represented in these datasets were obtained in the same manner, and included samples from RHS Hyde Hall 'Global Growth' garden in the UK.
A summary of the samples sequenced with the Illumina MiSeq is given in Table 2.1.

2.1.2. Survey Samples

Further tubers and plants were purchased from online vendors in the UK, Ireland, France, Poland and Hungary, via eBay and one commercial website (UK-based). These included one oca sample sold as stems for planting, one oca plant in a 1L pot labelled as 'Giggles', and a number of tubers. Ornamental plants from the genus *Oxalis* were also purchased. *O. palmifrons* and *Oxalis* 'Blue pea' arrived planted in soil. Other samples arrived as groups of bulbils. Leaf tissue from wild-growing *Oxalis spp*. was sampled from one site in Poole, England and three sites in York, England. The origins of the samples for the survey are summarised in Table 2.2.

2.1.3. Material from an Irish Commercial Nursery

Further tubers and plants were imported from a commercial nursery in Ireland. Tubers were subject to RNA extraction directly, while whole plants were sampled from leaf tissue (500mg). The tuber samples were collected together and pooled per variety, summarised in Table 2.3.

2.1.4. Accessions from the Royal Horticultural Society

Leaf samples were also obtained from material growing in the alpine collection at RHS Wisley. Plants were sampled randomly from leaflets, following the same protocol as mature oca plants. On arrival the leaflets were kept in storage at -80°C until RNA extraction and testing with qRT-PCR.

Table 2.1 A table showing the origins of the samples which had their total RNA extracted, which was then sequenced via HTS. Samples labelled as '2019' were analysed using HTS and Angua prior to the present work and data were provided by Ian Adams of Fera Science Ltd.

Sample	Variety	Source	Notes
PPN	Pink Perfection	Internet (Netherlands)	2019
NP1	Negra	Internet (Poland)	2019, same purchase as NP2
NfP	New from Peru	Internet (Poland)	2019
YfP1	Yellow from Peru	Internet (Poland)	2019, same purchase as YfP2
HHPP	Pink Perfection	RHS Hyde Hall (UK)	2019
FG	Unknown	Internet, free gift (Rothsay, UK)	2019
NP2	Negra	Internet (Poland)	2019, same purchase as NP1
YfP2	Yellow from Peru	Internet (Poland)	2019, same purchase as YfP1
LO	Long Orange	Internet (Poland)	2019
HHPP	Pink Perfection	RHS Hyde Hall (UK)	2019
HHU1	Unknown	RHS Hyde Hall (UK)	2019
HHU2	Unknown (Guild of Oca Breeders)	RHS Hyde Hall (UK)	2019
U1	Unknown (Guild of Oca Breeders)	Internet (Unknown)	2019
U2	Unknown (Guild of Oca Breeders)	Internet (Unknown)	2019
Nar	Narouhal	Internet (Unknown)	2019
EP	Elizabeth's Pink	Internet (Unknown)	2019
CY	Yellow	Internet (commercial supplier, UK)	2019
U3	Unknown	Internet (Unknown)	2019
R1	Red	Internet (Unknown)	2019
PL1 to **PL27**	Unknown	Internet (Poland)	Fungus, shipped in newspaper, labelled as 'flavourings'
DK1 + **DK2**	Dylan Keating's	Internet (France)	Good condition
OB	Orange Beauty	Internet (France)	Good condition
LY	Long Yellow	Internet (France)	Good condition
B1	Unknown	Internet (Poland)	Good condition
Amar	Amarillo	Internet (France)	Good condition
An	Anilla	Internet (France)	Good condition
OAEC	OAEC	Internet (France)	Good condition
WP	Will's Pink	Internet (France)	Good condition

Table 2.2 A table of the survey samples, their origin, how many were sampled per variety, and what tissue was sampled.

Species	Variety	Origin	Received	Sampled	No.	Notes
*O.palmifrons**	Unknown	UK	P	L	1	Plant passport, good condition
T. majus	Blanca	Ireland	T	L	1	No plant passport
O. tuberosa	White	Poland	T	T/L	3/3	No plant passport
Oxalis sp.	Blue pea	UK	P	L	1	No passport, yellowing and wilting
Oxalis sp.	Unknown	UK	B	B	5	No plant passport
O. corymbosa aureo-reticulata	Unknown	UK	B	L	5	Plant passport
O.versicolor	Candy cane sorrel	UK	B	L	1	No plant passport
*O. triangularis papillonacea**	Irish Mist	UK	B	L	1	No plant passport
O. triangularis	Unknown	UK	B	L	2	No plant passport
O. tuberosa	"Rainbow mix"	Ireland	T	L	5	No plant passport
O. tuberosa	Oca de Perou	UK	T	T/L	5/3	Came with soil
O. tuberosa	White Wonder	UK	T	T/L	3/3	No plant passport
O. tuberosa	"Red"	UK, com.	T	T	2	No plant passport
O. tuberosa	"Yellow"	UK, com.	T	T	2	No plant passport
O. tuberosa	"Cream"	UK	T	T/L	5/5	No plant passport
O. tuberosa	New Zealand Red	UK	T	T	1	Declared as 'cosmetics'
O. tuberosa	Unknown	UK	T	T	1	No plant passport
O. tuberosa	Unknown	UK	S	S	1	No plant passport
O. tuberosa	Giggles	UK**	P	L	1	Tuberising, no plant passport, mottling on leaves

*Came with a plant passport. **Variety registered by a nursery in Germany.
T = tuber, L = leaf, S = stem, P = plant in soil, B = bulb.

Table 2.3 A table of the varieties and tissue types provided by the commercial nursery in Ireland.

Species	Variety	Number	Tissue
Yácon	Dimi	1	Leaf
Yácon	Morado	1	Leaf
Yácon	New Zealand	1	Leaf
Yácon	Peru	1	Leaf
Ulluco	Colombia	5	Tuber
Mashua	Unknown	1	Leaf
Oca	Apricot Delight	5	Tuber
Oca	Halford's Red	5	Tuber
Oca	Morten Smith Red	5	Tuber
Oca	Mellow Yellow	5	Tuber

Table 2.4 A table of the *Oxalis sp.* samples provided by RHS. These samples are from specimens kept at RHS Wisley in the UK. Origins of the species are according to the RHS database, Gardner et al. 2003, or Germishuizen and Meyer 2003.

Species	Accession	Species Origin	From
O. gigantea	W20151548A	Chile	
O. brasiliensis	W883287A	S Brazil to NE Argentina	
O. triangularis	W20140655A	Andes, Brazil, SE South America	CN
Oxalis "Ute"	W20034433A	Unknown	CN
O. pedicaria	W20014786A	South America	PSC
O. melanostricta	W880075	South Africa (Cape)	
O. ambigua	W20014707	South Africa (N & W Cape)	CN
O. callosa	W20081712	South Africa	CN
O. pulchella var. tomentosa	W20081720	South Africa	CN
O. tetraphylla	W20191749*A	Mexico	
O. tetraphylla	W20191749*B	Central America and Mexico	CN
O. hirta 'Gothenburg'	W953100	South Africa	CN
O. obtusa amber-flowered	W20081729	South Africa	CN
O. purpurea cherry-flowered	W20081723	South Africa	CN

*CN = Commercial Nursery, PSC = Plant Sales Centre

2.1.5. Positive Controls

In order to test the specificity of the ONV1 primers, positive material was purchased which consisted of cherry leaf roll virus (CLRV) nucleotides sourced from DSMZ, Germany (PV-0797 and PV-0432), grapevine fanleaf virus (GFLV) nucleotides sourced from DSMZ (AS-0205), and positive controls (RNA eluted in molecular grade water) from Fera Science Ltd. of blackcurrant reversion virus, tobacco ringspot virus, tomato black ring virus, and tomato ringspot virus.

2.2. High Throughput Sequencing

2.2.1. Sampling

Tubers were planted on arrival in 1L pots using general purpose potting mix for sprouting. A representative sample of the mplants were kept based on available glasshouse space and possible symptoms. At least one plant of each labelled variety was kept, if they emerged.
The number of plants sampled, tissue types sampled for RNA extraction and final read numbers of each HTS run of internet-traded oca are summarised in Table 2.5.
All plants were transferred to 3l pots once they reached maturity. Lighting conditions were between 04:00 and 20:00 hours when ambient light dropped below 90Wm-2 and not exceeding 250Wm-2. Temperature was kept at 18°C at night and 22°C during the day.
Any plants with mottling of the leaves or deformed leaflets were chosen to continue growing on in case the cause was viral, along with a random selection of asymptomatic plants. The remaining living plants after harvesting were 28 ordered from Poland, 8 ordered from France, and 5 older specimens from 2019.
Thirty-two days after planting, leaf tissue was sampled randomly from the plants, prioritising any leaves with possible mottling, flecking, twisting etc. These leaves were then pooled, stacked

Table 2.5 A table of HTS datasets and which plants went into them. The table summarises the number of plants whose RNA was extracted to form the pool, the library type, whether or not an additional CTAB step was required, read numbers before and after trimming, and the tissue sampled.

Run	Sample	Pre-trimming	Post-trimming	Number	CTAB?	Library	Tissue
L15S20	**PPN**	1879336	1794265	1	Y	ScriptSeq	Tuber
L15S6	**NP1**	795230	795230	1	Y	ScriptSeq	Tuber
L15S7	**NfP**	1519235	1513541	1	Y	ScriptSeq	Tuber
L15S8	**YfP1**	57704	551810	1	Y	ScriptSeq	Tuber
VMS6	**HHPP**	1785348	1633353	1		TruSeq	Tuber
VMS18*	**HHPP**	1318648	1045815	1		TruSeq	Tuber
L17S14	**FG**	543047	47756	1	Y	ScriptSeq	Tuber
L17S17	**NP2**	652626	640876	1		ScriptSeq	Leaf
L17S18	**YfP2**	756692	725979	1		ScriptSeq	Leaf
L17S19	**LO**	887748	848204	1		ScriptSeq	Leaf
L20S11	**HHPP**	1879336	1794265	1		TruSeq	Tuber
L20S12	**HHU1**	2211	1157	1		TruSeq	Tuber
L20S13	**HHU2**	39146	15088	1		TruSeq	Tuber
L21S1	**U1**	741120	737678	1		TruSeq	Tuber
L21S2	**U2**	604533	599193	1		TruSeq	Tuber
L21S4	**Nar**	494566	489691	1		TruSeq	Tuber
L21S5	**EP**	688299	655799	1		TruSeq	Tuber
L22S3	**CY**	907423	903589	1		TruSeq	Tuber
L22S6	**U3**	1619235	1529375	1		TruSeq	Tuber
L23S3	**R1**	1035451	1026225	1		TruSeq	Tuber
IS-BLK	**Unknown**	208895	152896	5	Y	TruSeq	Leaf
PL-BLK	**PL1-27**	1358720	1235432	27	Y	TruSeq	Leaf
FR-BLK	**Mixed****	6349578	6040664	9	Y	TruSeq	Leaf
PL24	**PL24**	816722	778370	1		TruSeq	Leaf

*No ribo-depletion step. **Mixed pool of varieties: DK1, DK2, OB, Amar, An, OAEC, WP, LY, B1

together and a 5mm borer was used to punch out a stack of discs of tissue (approx. 500mg) for RNA extraction and HTS sequencing.

The nucleic acid extracts sequenced as sample PL-BLK were extracted from a pool of tissue from 27 of the plants originating from Poland. FR-BLK was a pool containing all 8 French plants and one Polish plant. The varieties for the French plants were: one plant of Orange Beauty, two plants of Dylan Keating's, one of Amarillo, one of Anilla, one of 'OAEC' Pink and one of Will's Pink. The sample IS-BLK consisted of pooled tissue from the five plants purchased in 2019. When sequencing the nucleotides in tissue from the commercial material from Ireland, total RNA was extracted from approx. 500mg of tissue taken from the tuber core. The yacón samples and mashua samples were sampled from leaf tissue in the same manner as individual oca plants.

2.2.2. Nucleic Acid Extraction

Extraction followed a modified CTAB protocol (Adams et al., 2009), including on-column DNAse digestion using an RNAse-free DNAse kit (Qiagen) when an RNEeasy Plant Mini Kit (Qiagen) was found to be unsufficiently high quality for HTS, determined by adding 1ul of

eluted nucleotides to a NanoDrop One (ThermoFisher). Samples extracted in this manner are indicated in Table 2.5. The protocol was modified to extend LiCl precipitation step was left overnight, approximately 12 hours.
When preparing samples for nucleotide extraction and high through sequencing, homogenisation of all samples utilised snap freezing with liquid nitrogen followed by grinding with a mortar and pestle.

2.2.3. Library Preparation

Library preparation was achieved using a dual-indexed TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina) or ScriptSeq Complete kit (Plant leaf) (Illumina) library kit according to manufacturer's instructions, utilising a 96 well ITS UDI dual-indexed plate for multiplexing. 0.1-1µg of purified RNA was used as an input. Final cDNA concentration of the individual and pooled samples was determined using a Qubit (ThermoFisher) fluorometer with a Qubit dsDNA (ThermoFisher) kit according to manufacturer's instructions. The length of the fragments in the pool were determined using a Tapestation (Agilent). When preparing the PL-BLK pool for sequencing, the Tapestation results indicated a 150bp peak, which was confirmed to be due to adapter dimers. Subsequently, the pool was cleaned again using Ampure XP (Beckman) beads with a ratio of 0.7x beads, which reduced the peak resulting from the primer dimers. The concentration of the pool was raised to sufficient levels using a final bead clean with a 2x ratio, leading to a 447 bp peak with a 8.80 nanomolar cDNA pool. No fragmentation step was necessary due to the small average size of the cDNAs in all instances. Each pooled sample formed one library for sequencing. Individual runs constituted 22 individual libraries including a negative healthy (*Nicotiana tabacum* cv. White Burley) control with ERCC (ThermoFisher, USA) spiked in (24 indexes total). The healthy control consisted of nucleic acids extracted from a plant maintained in a controlled environment kept free of viruses and other pests, grown from seed in the facility.

2.2.4. Sequencing and Bioinformatics

Sequencing was conducted on an Illumina MiSeq (Illumina) using a V3 reagent kit. The read length for the kit was 301 bp in both directions of the paired end read. These were all multiplexed on one lane, as the MiSeq does not have multiple lanes.
Initial analysis of HTS runs used the Angua3 pipeline (Fowkes et al., 2021b). Briefly, the demultiplexed reads were trimmed to q10 and minimum 50 bp length and the adapters removed with BBDuk (BBTools, B. Bushnell sourceforge.net/projects/bbmap/) or Sickle to Q10 (PHRED score of 10). These were assembled using Trinity (Haas et al., 2013). All the contigs of 200nt or greater were subject to a search with BLASTN against a complete local copy of the NCBI GenBank nt database (Camacho et al., 2009). Reads greater than 1000 bp in length were checked for homology to viruses using BLASTX against a local copy of the NCBI nr database. The output files of these searches were analysed, visualised and extracted using the metagenomic software, MEGAN community edition (Huson et al., 2016). Later HTS datasets

were analysed using a version of the Angua pipeline refactored by the author, which uses BBDuk and assembles the demultiplexed reads with either Trinity or SPAdes. This modified pipeline includes extensions by this author for ORF investigation and programmatic mapping of inputted reads to extracted contigs (expanded functionally initially implemented by Sam McGreig of Fera Science Ltd.), and is available as the conda environment Angua_Luggage and on github (https://github.com/mwodring/Angua_Luggage).

Demultiplexed reads from the PL-BLK dataset and the PL24 dataset were concatenated together before processing using the Angua3 pipeline to recover a more complete genome than either sample alone. BBMap was used to infer average coverage of contigs, by mapping trimmed reads back to the generated contigs. Analysis of the results outputted by BBMap used samtools to generate histograms and average coverage data from the mapped files (Li et al., 2009) and IGV to visualise the alignments (Robinson et al., 2011). Percentage identity between sequences was determined by aligning the nucleotide or protein accession of the closest related known virus to the genome or predicted protein respectively of the novel virus sequence using Geneious Prime (version 2019.2.3) (Dotmatics).

The RNAFold web server (Gruber et al., 2008) was utilised to predict secondary structure.

For each novel RNA virus candidate, where available, the putative replicase region was used to query the Serratus PalmID website (Edgar et al., 2022). Hits with greater than 90% identity to accessions in the NCBI database were extracted and aligned with the entire replicase sequence using MUSCLE. This is because Serratus PalmID only detects the 'palmprint', a small sequence representing the catalytic core of the RdRp. Where 90% homology was detected to reads from the SRA, these matches were examined and where appropriate, a selection of the SRA archive data representing the highest homology in reads from samples from different geographical regions were downloaded using SRAToolkit. These reads were trimmed with BBDuk and then mapped to the genome of the putative novel viruses using bwa-mem2. Reads were then assembled using Angua3, terminating the pipeline at the BLASTN step. Sequences of the corresponding genera were extracted (in case of mapping drop-off due to distant homology within the same species i.e. around 90% (Edgar et al., 2022)). These reads were then compared against the draft genome of the novel virus using BLASTN.

Data were also utilised from 2019, collected prior to the beginning of this PhD. These constituted a number of datasets generated from plants purchased via internet shopping. Outputted sequences, sorted per genus based on MEGAN assignment as well as demultiplexed reads were obtained from Ian Adams (Fera Science Ltd.). The sequences from these 2019 datasets were investigated to determine presence of viruses detected in the pooled datasets from 2021. In some cases the sequences generated from the 2019 datasets provided more coverage of the genome of the respective virus than the datasets sequenced in 2021. Datasets sequenced in 2019 and the samples from which the respective RNA was extracted are indicated in Table 2.5 on Page 31.

NCBI ORFfinder (Wheeler et al., 2007) and Geneious Prime were used to determine open reading frames and extract predicted protein sequences of putative novel viruses for alignment and comparison. These two resources were used in conjunction with one another to

cross-reference findings and be more confident in their interpretation. A BLASTP search of the amino acid sequence of predicted ORF products was also used for confirmation.
Determination of cleavage sites of the novel nepoviruses was achieved by reviewing the literature of nepoviruses and their 3C-like cysteine protease, the details of which are outlined in paragraph 3.2.11 and paragraph 3.2.11. In addition, comparisons were made using Geneious Prime between the 5' UTRs of both genomic RNAs, as well as the N-terminal region of the polyprotein ORF. InterProScan (Jones et al., 2014) was queried with predicted protein sequences to identify common motifs and transmembrane domains. This provided greater confidence in predicted cleavage sites.
Predicted proteins of the novel potyvirus were compared with proteins from other potato virus Y subgroup viruses using SDT (Muhire et al., 2014). The output of this comparison was visualised using the Python library Seaborn.

2.3. PCR

2.3.1. Primer Design

Primers were designed to detect each virus following guidance from Neil Boonham (pers comm) as well as general considerations for conventional PCR and quantitative PCR primer design (Dieffenbach et al., 1993, Bustin and Huggett, 2017).
ONV1 primers (RT-PCR and qRT-PCR) were designed using Primer Express 2 (Applied Biosystems); as these were for detection of the virus within the plants the virus was first detected in using HTS, the design process did not attempt to exclude related viruses using, for example, sequences sourced from GenBank during primer design. However, all primers designed for this study underwent a basic screen for non-target binding per the recommendation in the Geneious Prime guidelines, using NCBI GenBank as a reference. The exception is the qRT-PCR ophiovirus primers.
Primers are summarised in Table 2.6. qRT-PCR primers were designed for detection of viruses and RT-PCR primers were designed for some species in order to confirm the HTS results using Sanger sequencing.
Geneious Prime was used to design the remaining primers except Arabis mosaic virus (ArMV) using Primer3 (Untergasser et al., 2012). The primers for detection of ArMV were designed using sequence comparison and manual selection of degenerate individual bases, as computerised primer design was not able to produce primers which would not amplify grapevine fanleaf virus sequence while still detecting ArMV.
As testing with ELISA failed to produce a positive result for ArMV in the plants, primers were designed to detect the ArMV population sequenced from oca.
Other primers used are: cucumber mosaic virus qRT-PCR primers (Skelton et al., 2018), *Nepovirus* subgroup C conventional primers (Digiaro et al., 2007), ArMV 'Buffer A' qRT-PCR primers designed to the coat protein of a wide variety of isolates at Fera Science Ltd. (unpublished), *Potexvirus* generic conventional PCR primers (Van Der Vlugt and Berendsen,

2002), *Ophiovirus* generic conventional PCR primers (Vaira et al., 2003) and Ullucus polerovirus 1 qRT-PCR primers (Fox et al., 2019). The positive controls used were primers designed to detect the plant COX(Weller et al., 2000) gene and primers designed to NAD5 were compared as an alternative positive control (Botermans et al., 2013).

Table 2.6 A table of the primers designed to some of the novel viruses found in oca and to Arabis mosaic virus.

Virus	Name	Sequence	Type	Region
ONV1	ONV1-RNA1-F	CCCTATGAAATCACTAGACGAAGCT	qRT-PCR	Polymerase
	ONV1-RNA1-R	CACCTCGAAGTTCATGTTGCA	qRT-PCR	Polymerase
	ONV1-RNA1-Pe	TGGTGGAGGACATTCTAGAGCAGTGGC	qRT-PCR	Polymerase
	ONV1-RNA2-F	GGTTTTATGGGTTATACTGGTTGTTACC	RT-PCR	X4-MP
	ONV1-RNA2-R	TTCAATCGCGTCAATGGAGAT	RT-PCR	X4-MP
OVX	OVX-c-F	CTTGTGCGACGTCTAATGCG	RT-PCR	TGB3*
	OVX-c-R	CAAAGCCACGGAAAGACAGC	RT-PCR	TGB3*
	OVX-q-F	CGTTTGCCGTTGCGCTCAT	qRT-PCR	Replicase**
	OVX-q-R	CATACCACCGCCTCCAGGT	qRT-PCR	Replicase**
	OVX-q-Pe	CGCTCACTGGGTAGATCAGGTTTCCGTCG	qRT-PCR	Replicase**
ArMV	ArMV-F	GAGCTGCGGCACTCTTTG	qRT-PCR	CP
	ArMV-R	CATACCACCGCCTCCAGGT	qRT-PCR	CP
	ArMV-Pe	AGTACCCYCAAGGTGCTCCTTTGGTTGTCC	qRT-PCR	CP
OAV1	OAV-F	TGATGACTTGGTCCGTGACAA	qRT-PCR	Replicase
	OAV-R	AAAGTCGGGTTTCCAGGTTTCG	qRT-PCR	Replicase
	OAV-Pe	AGTTGCTGGAGATACTGCCAAGGCCAATTC	qRT-PCR	Replicase
OCV1	OCV1-F	CCATCTCTTGGGCCTTCTCT	qRT-PCR	V3/V4 (MP)
	OCV1-R	TAGTCGACGAATTGCCCAACA	qRT-PCR	V3/V4 (MP)
	OCV1-Pe	ACGGGAGACTACAAGCTTTTTGCACGCA	qRT-PCR	V3/V4 (MP)
Ophiovirus*	Oph-q-F	AGTTGACGCTGATGCTTTACTT	qRT-PCR	Replicase
	Oph-q-R	TTCAAGCAGAAGACGGCATAC	qRT-PCR	Replicase
	Oph-q-Pe	CGGAAGAGCACACGTCTGAACTCAGTCA	qRT-PCR	Replicase

* Sequencing primers only; these likely will not detect every sequence of the isolate they were designed to and are not recommended as detection primers for this reason. ** 'France' isolate only.

2.3.2. Nucleic Acid Extraction for PCR

For RT-PCR (reverse transcription PCR, or RT-PCR in this thesis) and qRT-PCR (quantitative reverse-transcription PCR), each individual plant constituted its own sample. Positive controls were identified through positive qPCR results and used individually for their respective assay as the study progressed. Initially, positive controls were created as pools of RNA extracts from the individual plants. Due to a lack of knowledge of which plants contained the virus initially, extracts from individual plants forming the pool in which the virus was detected were pooled together by combining 10ul of each extract into one microcentrifuge tube, without normalisation. These pools were as used for samples sequenced together as the initial HTS runs. As individual plants were found to be positive, ideally with Ct 30 or lower, those extracts were used as positive controls for later tests, including the survey. Nucleic acid extracts used as input to HTS library preparation were not utilised as positive controls, in order to confirm the presence of viruses from a new extract.

All RNA extraction used tissue comprising of sets of three whole trifoliate leaflets. For plants inoculated mechanically, tissue was harvested from new leaf growth or from symptomatic material. Each sample was approximately 0.3g grams of tissue.

Bulbs/seeds were soaked in thiocyanate buffer for one hour then crushed with a mallet before following the standard protocol outlined below.

A comparison was made between oca RNA extracted with a Kingfisher mL or Flex and Invitek beads (Fox et al., 2019, Fowkes et al., 2021b); RNA extracted with solely CTAB (Chang et al., 1993); and RNA extracted with an RNEasy Plant Mini Kit using the same method outlined below. RNA quantity (determined using a NanoDrop One) was found to be equivalent between the different extraction methods on the same oca samples. For this reason, the method below was chosen for PCR applications, as it is both cheap, easily scalable, and time efficient.

Extraction of RNA for qRT-PCR and RT-PCR amplification was conducted using a Kingfisher mL Purification System (Thermo Scientific) machine or a Kingfisher Flex System (Thermo Scientific) with magnetic bead extraction (Map A Solution, Invitek).

Samples, fresh or frozen at -80 c, were homogenised using a Homex grinder in a BioReba universal grinding bag with mesh before adding 1mls/0.1g tissue of quanidinium hydrochloride buffer with 2% TNaPP and 1% Antifoam B. A total of 1.5-2ml of the lysate was transferred to a microcentrifuge tube, using the sap opposite the mesh to exclude debris. The samples were centrifuged at 5.5 xg for 1 minute and 750 µl of the cleared lysate was recovered.

For extractions using a Kingfisher mL Purification System, cleared lysate was added to commercially available racks in the first well along with 40ul of Map A beads. The other wells contained in order: 1ml lysis buffer as above, 1ml 70% ethanol, a further 1 ml of 70% ethanol, followed by 200ul molecular grade water for elution of the RNA from the beads. For extraction using the Kingfisher Flex System, each deep well plate constituted a different step in the process rather than a different sample and elution was until 400ul of molecular grade water (MGW). The eluted nucleic acid extracts were stored at -20°C until testing. This method was also used for freeze-dried, lyophilised material from the 2019 study.

2.3.3. qRT-PCR

The cycling conditions for all qRT-PCR primers were as follows: 10 minutes at 50°c, followed by two minutes at 95°c, then 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. All reactions were performed on a QuantStudio 6, 7 or 12 instrument (ThermoFisher) using an iTaq Universal One-Step RT-qPCR Kit (Bio-Rad). Primer concentrations were 0.375 µl and probe concentrations were 0.25 µl.
All RT-PCR reactions terminated in an indefinite hold step at 20°C and had a final reaction volume of 25 µl.

Optimisation

Specificity ArMV and ONV1 qRT-PCR primers were tested against nucleic acids extracted from cherry leaf roll virus (CLRV) positive birch sourced from DSMZ, grapevine fanleaf virus (GFLV) sourced from DSMZ, and positive controls from Fera Science Ltd. for blackcurrant reversion virus, tobacco ringspot virus, tomato black ring virus, and tomato ringspot virus.

Inhibition test Due to suspected inhibition from secondary metabolites in the oca plants, inhibition was tested for using two methods. In the first test, one pair of wells was spiked with 1 µl of RNA extract from oca, and the other with 1 µl of a 1:10 dilution extract in molecular grade water (5 µl sample to 5 0µl MGW). Because dilution reduces inhibition (though may reduce sensitivity)(Hull, 2001), this was undertaken to account for potential inhibition with new primers and reaction conditions.
Further PCR reactions (see subsubsection 2.3.4) were performed using neat extract only without dilution. It was therefore presumed that any high Ct values were due to poor primer design and/or low viral titre in the plant sample.
The second test for inhibition involved spiking oca extracts and*N. tabacum* with potato spindle tuber viroid. The Ct values for both *N. tabacum* and oca when tested for this viroid were comparable, making it unlikely that compounds present oca tissue interfere with qRT-PCR as performed here.
The threshold for a positive was a Ct value of 38 or less for both technical replicates. Any sample with a Ct value between 38 and 40, or where only one of the duplicate wells produced a positive result, were tested again for confirmation. If, on confirmatory testing, both replicates produced a Ct value between 38 and 40 the result was considered to be weakly positive. If only one replicate produced a positive result, this was considered as contamination and the test was considered to have produced a negative result. Confirmatory testing used the same extracts as the original test, due to lack of availability of additional sample material in most cases.

Positive controls All samples were tested for successful extraction using COX primers (Weller, 2000). Though COX is present in DNA form as it is part of the plant gene, the RT-qRT-PCR mastermix was used for consistency. Primers designed by Botermans et al. to the NAD5 gene were compared as a possible positive control which tests directly for RNA

(Botermans et al. 2013). A secondary test of the NAD5 primers involved a BLASTN search using the Botermans et al. primers against all sequences in the dataset derived from HTS of RNA extracted from the pooled sample PL-BLK.
For this reason the COX positive control was utilised as a control for successful extraction, and inhibition (e.g. *Chenopodium quinoa* has consistently higher Ct values for this test than does *N. tabacum*).

2.3.4. RT-PCR

All RT-PCR reaction products were visualised using 1 µl of ethidium bromide per 100 ml of liquid agarose in 1-2% agarose gel with a 1 kb ladder (New England Biolabs). The bands were illuminated and visualised using a DigiDoc-It Transilluminator (UVP). All Promega AccessQuick RT-PCR System and (Promega) ('Promega kit') and OneStep RT-PCR Kit (Qiagen) ('Qiagen kit') reactions used 0.4 µM forward and reverse primers. SuperScript IV One-Step RT-PCR System (Invitrogen) ('SuperScript kit') reactions used 0.5 µM of forward and reverse primers. All reactions were 25 µl with 1 µL of template.

Cycling Conditions

Nepovirus subgroup C As preparation for testing oca nepovirus 1 (ONV1), CLRV-positive material from DSMZ PV-0797 and PV-0432, was tested using primers developed for subgroup C nepoviruses, with a predicted 640 bp product (Digiaro et al., 2007). These primers were tested with three different commercial mastermixes with different cycling conditions:
Qiagen kit: 50°C for 30 minutes, 95°C for 15 minutes, then 35 cycles of 94°C for 1 minute, 50°C for 1 minute, 72°C for 50s. The last step was 72°C for 10 minutes.
SuperScript kit: 55°C for 10 minutes then 98°C for 2 minutes, followed by 35 cycles of 98°C for 10s, 65°C for 10s, 72°C for 1m48, and a final cycle of 72°C for 5 minutes.
Promega kit: 45°C for 45 minutes, 95°C for 2 minutes, then 35 cycles of 95°C for 1 minute, 50°C for 1 minute, then 72°C for 45s, followed by a cycle of 72°C for 5 minutes.

Oca nepovirus 1 The Qiagen OneStep RT-PCR kit was found to be the most effective at producing bands of the appropriate size (1235bp). Cycling conditions were: 50°C for 30 minutes, 95°C for 15 minutes, then 40 cycles of 94°C for 30s, 60°C for 1 minute, 72°C for 90s, then one cycle of 72°C for 10 minutes. The band at 1235 bp was extracted for sequencing. Tests with this primer used extracts that previously produced a positive result when tested with the qRT-PCR primers designed to detect ONV1.

Ophiovirus The expected band size is 136 bp (Vaira et al. 2003). Cycling conditions are: 50 celsius for 30 minutes, 95 celsius for 15 minutes, then thirty cycles of 94°C for 30s, 50°C for 30s, 72°C for 1 minute, then one cycle of 72°C for 10 minutes. The mastermix used for this primer set was from a Qiagen OneStep RT-PCR kit.

Oca virus X The expected band size for was 215 bp. Cycling conditions are the same as for the ophiovirus primers, but with 35 cycles. The expected band size for the generic primers from Van Der Vlugt et al. (2002) is 280 bp, using the same conditions.

Sanger Sequencing

Sanger sequencing was used to confirm certain genomes where the HTS-derived viral sequence was ambiguous or non-typical of the genus. In particular, samples that produced a positive result when tested for ONV1 were subjected to RT-PCR to amplify a region within ORF2 which was predicted to terminate the polyprotein midway through the chain with a stop codon. Four samples were tested: the pooled RNA extracted from all PL-BLK plants, and three samples of RNA extracted from simple plants which produced a positive result when tested for ONV1 using qRT-PCR. These samples were also tested at a 1:10 dilution in MGW. Samples that produced a positive result when tested with primers designed to detect OVX were also subject to RT-PCR amplification using primers designed to amplify the TGBP3 region. The RT-PCR product that produced a band at the correct length was extracted and sent for Sanger sequencing.
cDNA from the RT-PCR reactions was extracted from agarose gels using a QIAquick Gel Extraction Kit (Qiagen), after excising the bands from the gel with an x-tracta Gel Extractor (Promega). Manufacturer's protocols were followed, the modification of a storage period in the QG buffer of 62 hours at 5°C before proceeding. The extracted samples were sent to Eurofins (Germany) for Sanger sequencing. The resulting chromatograms were quality trimmed to a 1% error rate with Geneious Prime before alignment to the viral sequence to which the primers were designed.

Rolling Circle Amplification

For identifying if the oca capulavirus 1 (OCV1) sequence detected in the oca plants derived from an episomal or integrated form, rolling circle amplification was performed using both the gene-specific primers designed to detect OCV1 and random hexamer priming, derived from the method in (Bömer et al., 2019). A positive was determined by comparing the Ct values from a qRT-PCR test of the control (before) and amplified (after) extracts.

2.4. Short Read Archive Data Mining

In expand knowdledge of the ARTC virome, NCBI's short read archive (SRA) was mined for viral sequence data. Using the binomial and common names of each ARTC species, the archives were queried for archived RNA-Seq data. Only RNA-Seq runs were included. WGA/DNA studies were excluded as this study type does not reverse transcribe extracted RNA in order to detect RNA viruses.
The remaining runs are summarised in Table 6.1, and were investigated using the same methods as the other HTS datasets, except for a small number which could not assemble with Trinity due to the read arrangement. SPAdes (Bankevich et al., 2012) was used to assemble these within the

Table 2.7 A table of the substitution model selected by IQ-TREE for each tree built, when restricted to viral models only. The region and molecule type are shown. All amino acid trees used MUSCLE for alignment and all nucleic acid trees used ClustalW for alignment.

Virus(es)	Model	Region (sequence type)
ONV1, ONV2	rtREV+F+R5	ProPol conserved region (aa)
ONV1, ONV2	rtREV+F+I+R4	CP subunits (aa)
OVX	rtREV+F+R5	CP (aa)
OVX	rtREV+F+I+R7	Replicase (aa)
Ophioviruses	rtREV+F+I+R5	CP (aa)
Ophioviruses	rtREV+F+R5	MP (aa)
OACRV1	rtREV+F+I+R9	Replicase (aa)
OCV1	rtREV+F+I+G4	CP (aa)
OCV1	rtREV+F+G4	Replicase (aa)
OCaV1	rtREV+F+I+G4	CP (aa)
OCaV1	rtREV+F+I+G4	RepA (aa)
OAV1	rtREV+F+G4	CP (aa)
OAV1	rtREV+F+R4	Replicase (aa)

modified Angua pipeline. All virus hits were extracted using MEGAN and further investigated individually. To obtain genome coverage and identities, extracted reads were mapped to a reference sequence from NCBI using the 'Map Reads to Reference' function in Geneious Prime. The reference sequence was chosen based on the BLASTN findings i.e. the accession with the highest identity to the contigs. For the turnip mosaic contigs from SRR2960161, reads which were clearly chimeric (i.e. originated from a fusion of host and viral sequence) were detected via BLASTN of mis-matching areas (usually to *Lepidium sp*.) and the entire contig was removed from the resulting assemblies.
In the case of SRR11431599, the reads were not equal between the two files and repair.sh from the BBDuk toolkit was used to amend this.

2.5. Phylogeny

Alignments were performed using MUSCLE using default settings (Edgar, 2004) for amino acids and ClustalW using default settings for nucleotides (Larkin et al., 2007) in Geneious Prime. Pairwise identity and identical sites were calculated with Geneious Prime. Trees were constructed in IQ-TREE (Nguyen et al., 2015) with 1000 bootstraps using UFBoot (Hoang et al., 2017) and the recommended SH-like approximate likelihood ratio test (SH-aLRT) (Guindon et al., 2010) the best fit model restricting to viral substitution types. Trees were then collapsed to only branches with 80/95 SH-aLRT/UF support using IQ-TREE. The resulting high-confidence trees were visualised with ITol (Letunic and Bork, 2021).
The phylogeny trees use the virus rather than species (binomial) names due to comparing a *contig representing* a putative novel virus species to *type specimens of* a virus.

Nepoviruses The phylogenetic analysis of RNA1 of the novel nepoviruses (with arracacha virus B accession YP_010840567, *Cheravirus*, outgroup) used the translation of the conserved

Pro-Pol region from the CG/SG (Pro) to the YGDD (Pol) motif used for species demarcation in *Secoviridae* (Sanfaçon et al., 2009). All three units of the CP region were selected for RNA2 as this region contains no motifs conserved throughout the family (Le Gall et al., 2008). Every available RNA1 polyprotein available as accessions listed in the ICTV database for members of the genus *Nepovirus* was utilised. Some exemplar isolates lacked complete CP or ProPol sequence on ICTV, and were excluded from the relevant analyses.
The potato virus B RNA1 and RNA2 ORFs were obtained via translation and comparison in Geneious Prime, as no entry in the protein database exists for the type specimen. Potato virus U accession AYF57455.1 and blueberry latent spherical virus accession BAL04700.1 had no CG motif in the Pro region sequence, but did have an SG, which gave a ProPol region of the correct size when extracted and can be found instead of CG in some nepoviruses (Sanfaçon, 2022). In the case of blueberry latent spherical virus accession BAL04700.1, there were several SG residues, and the correct one was chosen by alignment (residue 1389). The peach rosette mosaic virus accession AQT41504.1 lacked this region but AAB69867 did not, so this sequence was chosen instead of the exemplar for this tree. The sequences available for green Sichuan pepper nepovirus did not match any protein regions on Interpro, but had a CG-YGDD inter-region which was of the correct length and so was included.
In cases where the NCBI accession was not annotated, the capsid region was extracted using InterPro scan of the protein accession, between the start of the 'Nepo coat N' and the end of the 'Nepo coat C' region. It is worth noting that this used the domain hits, not the cleavage sites, which differ per species and are not always correctly annotated on the NCBI database (Sanfaçon, 2022). Since arracacha virus B accession QVX32641.1 matched no coat protein accessions on InterProScan, the three combined chains of the annotated accession of cherry rasp leaf virus (CAF21714) were used as an outgroup instead for this tree.

Poleroviruses For the polerovirus trees, all poleroviruses for which complete genomes are available were included. Because beet mild yellowing virus accession X83110 did not have an annotated P1-P2 accession, where the P1-P2 fusion protein on the exemplar accession or associated RefSeq entry (accession NC_003491), an annotated accession with >99% identity to the exemplar accession was chosen instead for that virus (CP and RdRp, accession KC121026.1). For maize yellow mosaic virus, only the accession from the Yunnan isolate, KU248489 was used and for cereal yellow dwarf virus only isolate RPS accession AF235168 was used. For pepper vein yellows virus 2 HM439608, sugarcane yellow leaf virus accession AF157029, ullucus polerovirus 1 accession MH978189 and potato leafroll virus accession D13954 (for which no isolate >99% identical to the exemplar was available with annotations) the putative P1P2 protein was determined using KnotInFrame. For PLRV also, the CP was determined using find ORFs in Geneious Prime. Ullucus polerovirus 1 was also included, both oca (L15S20) and ulluco isolates, as these were found in ARTCs, as well as both the turnip yellows virus accession X13063 and brassica yellows virus accession HQ388348 isolates of turnip yellows virus. The enamovirus pea enation mosaic virus 1 accession ADO86940 was used as an outgroup for both trees.

Ophioviruses For the putative novel ophioviruses, phylogeny used the complete coat protein sequence or the complete movement protein sequence of all exemplar isolates from the ICTV metadata resource and all ophioviruses reported from (Debat et al., 2023). In some cases either the CP or MP was incomplete or absent, and these species were excluded from the analysis for this reason. Waitzia ophiovirus was named as such in the paper, but the accessions listed were labelled on NCBI as Viola ophiovirus (accessions BK062738 and BK062739). For this reason it was excluded to avoid confusion. A search for boraginaceae-associated ophiovirus accessions as listed in the paper returned no results (accessions BK062663 and BK062664). As an outgroup, the ssRNA cytorhabdovirus alfalfa dwarf virus (accession KP205452) was chosen for both trees.

Oca virus X For placing oca virus X within the genus, the RdRp (ORF1) region and CP (ORF5) region were used. Clover yellow mosaic virus accession CYMRNA (D29630) was only available as a genomic sequence and had no putative ORF annotations, and so the likely RdRp and CP regions were extracted using Geneious Prime. Lagenaria mild mosaic virus accession AB546335 was excluded from the RdRp analysis due to only partial sequence being present. The unclassified viruses, Caladium virus X, Paris polyphylla virus X, potexvirus 1 LSD-2014, potexvirus ST4, potexvirus ST5, yam potexvirus 1 and yam potexvirus 2 were did not have complete genomes available and were excluded from both analyses. The unclassified potexvirus Opuntia virus was excluded from the CP analysis as this region of the genome is absent from NCBI. Cardamom virus X isolate Kerala accession DAZ85788 (Sidharthan et al., 2021a), which is not yet classified, was included due to its relatively close relation to oca potexvirus 1.

Oca capulavirus 1 For oca capulavirus 1, a comparison was made between the genome of all geminivirids with a complete exemplar genome, and the contig considered to represent the genome of the novel capulavirus. The novel capulavirus was compared to all four ratified exemplar sequences as well as a *Trifolium repens* isolate of trifolium virus 1 accession MW698818.1 (Ma et al., 2021), proposed as a capulavirus. Representative accessions from members of the closely related genera *Mastrevirus*, *Grablovirus* and sequences of mulberry crinkle-associated virus were included from Varsani et al. 2017 as well due to unusual genome organisation (Varsani et al., 2017). Tomato yellow leaf curl virus accession CAA33688, in the same family but less closely related, was used as an outgroup (Varsani et al., 2017).

Tropaeolum potyvirus 3 For Tropaeolum potyvirus 3, the nucleotide sequence and amino acid sequences of the main ORFs were used, comparing it to other members of the potato virus Y subgroup. Alstroemeria mosaic virus accession AB158522, Cotyledon virus Y accession JN572103, Amaranthus leaf mottle virus accession AJ580095, Amazon lily mosaic virus accession AB158523, Alternanthera mosaic virus accession EF442668 and Pfaffia mosaic virus accession AY485276 were excluded as their sequences are incomplete. For individual cleavage products, in cases where the exemplar isolate was not annotated sufficiently, a relevant protein accession of the same species was used and the accessions are visualised on the resulting trees. The amino acid and nucleotide sequence of the entire ORF was used to infer a phylogeny of this

virus within the potato virus Y group, which includes the above and some from the wider group (Duarte et al., 2014, Quenouille et al., 2013).
Pepper yellow mosaic virus accession AB541985 had no associated NIb protein accessions on NCBI and so the cleavage site was determined by alignment with pepper mottle virus accession M96425 (Q/G on the N terminal and Q/A on the C terminus for a 519 long mature product). Mashua virus Y accession MH680824 was compared with verbena virus Y accession EU564817 and the NIb peptide was extracted (Q/S, Q/A, 519 aa). Bidens mosaic virus accession KF649336 was compared with sunflower ring blotch virus accession KX856009 (Q/S, Q/V, 519 aa).
For the coat protein, wild potato mosaic virus accession accession AJ437279 had no complete coat protein accessions on GenBank and was compared with Peru tomato mosaic virus accession AJ437280 (R/A, 272 aa). Pepper yellow mosaic accession AB541985 was compared with pepper mottle (Q/A, 278 aa). Mashua virus Y was compared with verbena virus Y (Q/A, 272 aa). Bidens mosaic virus was compared with sunflower ring blotch virus (Q/V, 267 aa).

Oca alphacytorhabdovirus 1 The oca alphacytorhabdovirus 1 replicase was compared with the *complete* replicase of the exemplar isolate of all ratified and unclassified members of the genus *Cytorhabdovirus*, where available. The outgroup used was Sonchus yellow net virus accession AAA50385, a member of a different genus within the same family. The replicase of northern cereal virus accession AB030277 was extracted via find ORFs in Geneious Prime. A further, larger tree included all proposed *Alphacytorhabdovirus* member replicases first detected in a data mining study (Bejerman et al., 2023).

Oca caulimovirus 1 Similarly, the phylogenetic tree for oca calimovirus 1 excludes thistle mottle virus, which has no GenBank entry. The type solendovirus tobacco vein clearing clearing virus accession AF190123_3 was used as an outgroup because its coat protein and replicase proteins are not fused, as with caulimoviruses, which is unlike some other caulimovirids (Teycheney et al., 2020).

Potato yellowing virus Placement of the PYV isolates within the tree used all PYV isolates from Silvestre et al. 2020, comparing amino acid sequence of the relevant ORF (Mtr-Hel (1a, RNA1), RdRp (2a, RNA2), CP (RNA3); isolates where a gene was only partial were not included (Silvestre et al., 2020).

2.6. Mechanical Inoculation

Host species and varieties The species and cultivars used were *Cucumus sativa* var. Telegraph, *N. tabacum* var. White Burley, *N. benthamiana*, *N. clevelandii*, *N. occidentalis, N. glutinosa, Chenopodium quinoa*, *C. amaranticolor, Pisum sativum* var. Onward, *Phaseolus vulgaris* var. The Prince, *Gomphrena globosa*, *Brassica rapa* var. Yuki, *Beta vulgaris* var. Roberta, *Datura stramonium*, *Matthiola incana* var. Tudor Tapestry, *Solanum lycopersicum* var. Alisa Craig, V*erbena* × *hybrida* var. Bordeaux and *Verbena sp, Petunia* × *hybrida.*

Material retained from the 2019 study A small number of samples retained from the 2019 study were used for attempting mechanical inoculation onto a range of experimental hosts. Each sample was stored as a tuber at -80°C and processed the same way as the other tissue used for mechanical inoculation. A sample from an unknown variety, YfP1, YfP2, NP2, HHPP, FG and LO were retained and used for mechanical inoculation from thawed tubers.
Positive controls were frozen sap of *C sativa* inoculated with CMV, frozen sap of *Chenopodium quinoa* inoculated with ArMV, *Impatiens sp.* and frozen sap of *Nicotiana rustica* inoculated with tomato spotted wilt virus and sap extracted fresh from bean plants infected with broad bean wilt virus 1 maintained by serial passage. These were inoculated onto the respective plant species.

Material from 2021 One plant determined to be positive by qRT-PCR for ONV1 (unknown variety) was used for the initial testing (this plant was later determined to be infected with ArMV, PL24). Later, one *C. sativa* plant that tested positive was re-sampled and used to inoculate a further test range. For the later sapping onto *C. sativa*, a different plant negative for all other viruses, including ArMV, was used to inoculate the plants (PL27).
For the potexvirus assay, one plant which produced a positive result when tested using the generic primers was used, of the variety Dylan's Keating (DK1). Plants were tested for OVX and ArMV using qRT-PCR and CMV and ArMV using ELISA.
For inoculation, fresh material from a plant that tested positive for the given virus was used, or else frozen tuber left from the previous study.
For the fresh plants, a small portion of leaf material was removed from the plant, comprising typically six sets of six leaflets, and ground with approximately 2ml of fresh phosphate buffer in a chilled mortar and pestle. Buffer was prepared using 5.676 66.64mM di-sodium orthophosphate and 3.628g 66.65mM potassium di-hydrogen orthophosphate in 1000ml of distilled water, adjusted to pH7. The resulting slurry was rubbed onto the two cotyledons of *Cucumis sativa* plants, or the first true leaf of other plants tested. Leaf samples were taken at 14-16 days and 21-23 days post inoculation (dpi). Each sample weighed approximately 0.3g. The tissue was then frozen in a BioReba bag until further processing. Generally, the sampled tested at 21-23 dpi were tested, with the 14-16 dpi samples used to determine the onset of infection if a positive was found. Asymptomatic plants were tested using the PCR primers detailed above to detect infection.
One group of six *C. sativa* plants was inoculated after growing in the dark for 24 hours pre-inoculation. Tissue from a cutting of a nepovirus positive oca plant was used for inoculation. This dark period has been found to increase the effectiveness of inoculation before and was tested for this reason (Kassanis, 1957).
Positive controls for all inoculations with material purchased in 2021 were a *C. sativa* plant inoculated with CMV-infected sap. The negative control was a plant of the species being challenged with the virus which was not inoculated. All plants were inoculated at the two cotyledon stage before the emergence of the first true leaf.

Statistical test As one cucumber plant out of 30 inoculations was positive, a further test was devised to determine if this one positive was likely to be representative. The R script used to infer the binomial distribution was provided by Roy Macarthur and is included in Appendix A. It was determined that 107 plants would be needed to determine whether or not the one positive observation reflected unlikely circumstances with 5% confidence. 110 plants were tested ten at a time over the course of Spring and Summer of 2023. In each batch one negative control and one CMV+ positive control was used (22 additional plants). Samples were taken as above from new growth (sink) material and stored at -80°C until extraction.

A script for a statistical test on the first round of mechanical inoculations was provided by Roy MacArthur (Appendix A). This used a beta binomial distribution to determine the underlying baseline of one observation in thirty, as initially observed. Using this, a further 110 samples were inoculated over ten weeks and harvested at two weeks post-inoculation to determine if the two sets of observations were concurrent or if the original set may ave had differing baseline assumptions than the second (i.e. that the positive cucumber detection was an outlier and not representative of the host-virus relationship).

Table 2.8 A table of frozen tubers retained from the 2019 study and the plants they were inoculated onto.

HTS Sample	Variety	Viruses	Cs	Nt	Nb	Nc	Cq	Ca
L17S8	Yellow from Peru	ONV2, UlPolV1, OACV1, Ophioviruses	7	0	8	6	8	0
L15S8	Yellow from Peru	ONV2, OVX, UlPolV1, Ophiovirus	3	0	3	3	0	0
L20S3	Unknown	ONV2	3	0	3	3	3	0
L17S19	Long Orange	OVX, OACC1, UlPolV1	7	1	8	3	8	3
L20S11	Pink Perfection	Ophioviruses, OVX, OAV1, OACV1	3	0	3	3	3	0
L17S14	“Free gift”	OVX, OcaV1, OAV1, OACV1, ArMV, UlPolV1	3	0	3	0	3	3
L17S17	Negra	UlPolV1	3	0	3	3	3	0

Cs = *Cucumis sativa*, Nb = *Nicotiana benthamiana*, Nc = *N. clevelandii*, Nt = *N. tabacum,* Ca = *Chenopodium amaranticolor*, Cq = *C. quinoa*

Table 2.9 A table of inoculations attempted from different samples purchased during this study. Original sample refers to inoculating from oca leaf material infected with a given virus. From *C. sativa* refers to the inoculation of ONV1 from a positive *C. sativa* plant. Dark refers to the test where plants were placed in darkness before inoculation. Statistical test refers to the inoculation of *C. sativa* plants from fresh oca leaf over the course of several months. Data points indicate number of samples tested positive / number of samples challenged with the vius.

Test	Virus	Cs	Nb	Nt	No	Ng	Gg	Ds	ChQ	ChM	Pv	Ps	Br	Bv	Sl	Pxh	Mi	Ver
Original sample	ONV1	1/24																
From C. sativa	ONV1	0/2	0/2		0/2	0/2	0/2	0/2		0/2		0/2	0/2	0/2	0/2	0/2	0/2	0/2
Dark	ONV1	0/5																
Original sample	OVX	0/4	0/3	0/2					0/3									
Original sample	OAV1		0/2	0/2		0/2				0/2	0/2	0/2						
Statistical test		0/110																

Cs = *Cucumis sativa*, Nb = *Nicotiana benthamiana*, Nt = *N. tabacum*, No = *N. occidentalis*, Ng = *N. glutinosa*, Cq = *Chenopodium quinoa*, Cm = *C. murale*, Gg = *Gomphrena globosa*, Ds = *Datura stramonium*, Pv = *Phaseolus vulgaris*, Ps = *Pisum sativum*, Br = *Brassica rapa*, Bv = *Beta vulgaris*, Sl = *Solanum lycopersicum*, Pxh = *Petunia × hybrida*, Mi = *Matthiola incana*, Ver = *Verbena × hybrida.*

2.7. ELISA

ELISA (Clark and Adams, 1977) was used to test for CMV (Agdia) infection in the individual plants forming the pooled HTS samples, to test for serological cross-reaction between ONV1 and CLRV (strain ch and strain e, from Bioreba), and to test for ArMV (DSMZ) infection in the plants.

All antisera were diluted according to manufacturer's instructions. Samples were tested in duplicate along with a negative control of fresh healthy sap from *N. tabacum* and sap from infected plants which had previously tested positive or were purchased as positive controls. This sap was either stored at -80°C or extracted fresh from maintained stocks. A sample was considered positive if the absorbance was 3x greater than the average absorbance values of the healthy control.

Chapter 3. Virome of European Oca

3.1. Introduction

Prioritisation of risks and when to manage them is an essential part of biosecurity. In plant health, a 'no risk' approach to biosecurity is not possible; there are always trade-offs between risk and reward in trade (Macdiarmid et al., 2013). First published in 2018, the UK Biosecurity Strategy was updated in 2023, focusing, perhaps unsurprisingly, on zoonoses and human disease due to COVID-19 (Cabinet Office, 2023). This strategy focuses on a need for baseline surveillance of pathogens, including plant pathogens, data sharing, and a global plant pest alert system. While the updated UK Biosecurity Strategy document makes reference to trade, risk management and SPS agreements, it makes no mention of internet trade and the informal economy beyond a small mention of fast parcel couriers.

In recent years, a number of studies have leveraged HTS and viromics to baseline natural and managed ecosystems as well as their interface, including results from a collaborative network, INEXTVIR. These include surveys of pea crops, *Prunus spp.*, carrots and their wild relatives, tomato, melons and their weeds, and wild Iberian ecosystems (Fox et al., 2022, Fowkes et al., 2021b, Schönegger et al., 2023a, Khalili et al., 2023, Maachi et al., 2022, Temple et al., 2023, Zamfir et al., 2023).

In addition, great progress has been made in bringing HTS from solely research use into the realm of front-line NPPO diagnostics, such as a favourable comparison between HTS and other methods for post-entry quarantine in Australia (Gauthier et al., 2022) and guidance on validating HTS for diagnostics (Lebas et al., 2022). High throughput sequencing can be used in support of diagnostics, investigating diseases of unknown aetiology or forming part of a generic diagnostic suite (Adams et al., 2018d). Virus discovery as a result of this, however, often lacks follow-up studies to provide additional biological data beyond the original finding context (Rivarez et al., 2021, Massart et al., 2017). In particular, recommended processes for connecting sequences detected via HTS surveys to risk include 'large scale field surveys' to establish a viral baseline in the importing country (Macdiarmid et al., 2013, Olmos et al., 2018, Massart et al., 2017, Fontdevila Pareta et al., 2023). The present chapter includes a pilot study for surveying ARTCs purchased via the internet, utilising primers designed to HTS findings in order to investigate the prevalence of the novel viruses within oca plants available for purchase on the internet and within a small number of individuals of the genus already present in the United Kingdom, wild or managed. The primers alone do not fulfill the need to detect and distinguish circulating variants as the (Massart et al., 2017) as these primers are for detection only, rather than sequencing. However, PCR diagnostics are fast, cheap, and aid in understanding the present

geographical range of the novel viruses within the hosts specified. Specifically, such an endeavour helps build a portfolio upon which risk management can be based, and utilises a fairly straightforward, scalable method.

One method of managing risk in trade and plant health, particularly in clonal crops, is the existence of a formal seed system, whereby seeds for planting are procured from commercial companies or foundations which have been deemed, by the actor in question, to be free of relevant plant pests (Forbes et al., 2020). The seed system ensures not only freedom from regulated pests but also uniformity of phenotype and genetics, and in Europe all imported seeds must follow equivalent regulations to its own (Misra et al., 2023). In some cases, for example in low-income countries or for small-scale growers of heritage breeds, the informal seed sector may fill a gap where the formal seed sector cannot meet growers' needs (Forbes et al., 2020). Formal seed systems tend to provide stocks free of disease, meaning that crops grown from certified seed tend to have higher yields, but formal systems may price individuals out of the market or be excessively bureaucratic (Buddenhagen et al., 2017). Lack of access to sufficiently well-tested potato seed is responsible for a yield gap in sub-Saharan Africa of up to 50% (Kreuze et al., 2020a, Harahagazwe et al., 2018). Informal seed systems can still be effective in reducing disease load in clonal crops but only if diseases are properly identified and removed, which is difficult if diseases are cryptic or symptom expression varies with climate (Thomas-Sharma et al., 2017).

In the Andes where they originate, informal potato seed systems are ancient, common, and still effective in management of viral diseases in potato through a mix of genetic diversity, the effects of altitude, and land management (Forbes et al., 2020). Informal seed sharing of cowpea was found to be potentially *less* damaging due to the shorter range of exchange than international seed companies, and informal seed had fewer bacterial and fungal pathogens than formal seed (viruses were not tested for) (Biemond et al., 2013). International seed companies sometimes use favourable, warm climates to bulk up seed before export, leading to accumulation of virus (e.g. aphid-borne) that may then be exported (Jones, 2021). That is one hypothesis for the observed difference in seed quality in the cowpea study between foundation/commercial and shared seed (Biemond et al., 2013). Generally, a seed system that does not adequately screen the seed it supplies for pests can lead to pathogen spread (Forbes et al., 2020).

Long-distance, informal exchange of non-certified propagative material, then, can be seen to combine aspects of both. Allotment growers in the UK demonstrate high rates of seed sharing and seed recycling, partially due to lack of commercial seed from which to grow the crops they desire, which are often connected with the cultural background of the grower (Cunningham et al., 2018). In Europe, seeds must be of a variety registered to an official catalogue in order to be sold legally, for example, with certification requiring and testing for a consistency of phenotype indicative of the variety (Misra et al., 2023). An EPPO A1 quarantine pathogen has already been detected in a UK allotment as a result of what is likely to be internet trade (Fox et al., 2019). At least one of the plants sampled in the present work belongs to a legally registered variety, 'Giggles', sold from a small UK garden centre.

Figure 3.1 Photographs taken at Fera Science Ltd. of oca tubers from the 2019 internet purchases. A shows the shape and colour of specimens of 'Long Yellow'. B is the morphology and colour of 'Amarillo'. D is the most common type, a pink, round tuber with white shading in parts. C shows tubers infected with an unidentified folder. These are not the specimens mentioned from a later purchase, but the symptoms are the same.

This present chapter acts as another case study for how HTS can be used in support of plant health. The aim was to explore the virome of plants purchased via the internet trading pathway to investigate the potential risk that these tubers may pose, particularly from viruses which would not be detected via hitherto-known targeted assays, such as ELISA. This chapter covers the early, exploratory stages of risk assessment, identifying viruses associated with imported tubers. By purchasing individual samples via the same route that a typical consumer would, it allowed a direct assessment of some of the viruses likely to enter/circulate within the UK via this route, on oca plants.

3.2. Results and Discussion

Due to the nature of the investigation, which utilised literature to infer putative genomic arrangements, the results in this chapter are presented combined with the discussion.

3.2.1. Tuber Conditions and Morphology

The tubers sequenced in this chapter and their origin are summarised in Table 2.1 on Page 28 for HTS samples and Table 2.2 on Page 29 for the survey samples.
Of the oca plants, all those sold simply as 'oca' with no variety or colour declared had the same pink-red skin and pale yellow to white flesh, with a generally rounded or oval shape - notably, the plant labelled as New Zealand red shared this phenotype, and may suggest an Australasian origin for plants with this shape and colour. Generally, these plants are pink all over, but have paler, yellow or white patches, especially on the raised skin around the eyes.
Oca de perou had deep red skin and purple flesh, and were longer. Dylan Keating's resembles the common pink phenotype but tended towards a more yellow-orange skin. Long Plants

Figure 3.2 Plants sent as 'Oca de Perou' from the United Kingdom. They were received with growing instructions including advice to save seed between growing generations. They were unwashed and coated with soil. These plants had a dark purple skin once washed and were the only variety with purple interior flesh.

labelled as 'Long Yellow' are pale yellow in colour and as the name suggests, somewhat longer in shape. The 'cream' plants resembled these more than they did the white wonder plants, especially with their long shape, but had a white skin and yellow-white flesh like 'white' and 'white wonder'. White wonder are cream to white in colour, very uniformly pigmented, and very round, as were plants ordered as 'white'. Apricot Delight are a pale orange with yellow flesh. Anilla were yellow in colour, brighter than 'long yellow', as were Mellow Yellow. Halford's Red and Morten Smith's Red were similar to the common pink phenotype but a slightly darker red in the skin. Amarillo were bright yellow and round in appearance.
All oca plants had similar morphology above ground: bright green upper leaves, usually quite hairy, with purple lower leaves and thick, red stems. All share a crawling growth habit, quickly growing off the sides of their pots.
The quality of the tubers varied considerably. A number of tubers came washed, packaged in airtight bags, and clearly labeled per variety. Some contained care instructions and harvesting suggestions. A parcel of samples purchased from a seller in Poland (sampled from leaf material as PL-BLK) arrived wrapped in damp newspaper. Many of the tubers showed signs of a fungal infection due to either poor quality tubers or poor storage conditions. This parcel came labelled as 'flavourings'. Later samples also came with what is likely to be intentional labelling errors to evade customs inspection at the border; of particular note were one package declared as 'cosmetics'. One package came unwashed, with original soil. This choice was declared on the eBay listing as intentional 'for freshness'. In the case of arracacha, washing may introduce small lesions that contribute to bacterial decay, so there is some basis for this decision (Hermann, 1997), but inclusion of soil with plants for planting presents a high plant health risk (Evans et al., 2010).

3.2.2. Ascertaining sample origin

'European' oca is itself a conglomerate of South American, Australasian, North American and European varieties, adapted to various day-lengths and climates by farmers and breeders.
For example, Halford (of 'Halford's Red') is the family name of a grower in Aoteoroa-New Zealand who claims that his grandfather brought the original oca lines from Chile in the mid 19th century, and claims to be the person who popularised the vegetable commercially 100 years later (Vietmeyer, 1991).
Dylan Keating, as a variety, originates with a person of that same name who grew the original seed tubers in Cornwall, UK, while 'New Zealand Red' are likely derived from a supermarket purchase in Aoteoroa-New Zealand later bulked up in Surrey, UK (https://www.realseeds.co.uk/unusualtubers.html). The Dylan Keatings tubers for this study were purchased from France, suggesting that the material has crossed the border at least twice. Plants labelled as 'OAEC Pink' were also purchased for this study; Occidental Arts and Ecology Center (OAEC) is a facility in Occidental, California, which sells 'deep yellow', 'pink-blushed' and 'orange' oca plants (https://nursery.oaec.org/). Amarillo may be one of several varieties, most of which are either Andean or North American

(https://nursery.oaec.org/https://www.cultivariable.com/product/oca/oca-varieties/oca-amarillo/).

In one case the variety (Dylan Keatings, Cornwall, UK) samples crossed the France-UK border at least twice, once to reach the seller in France, and once to return to England when sold from France. These plants are thus demonstrably subject to an informal, in some cases illegal trade across borders, and many are not free of viruses.

From the sequences derived through HTS, more sequences belonged to novel viruses than belonged to viruses which were described prior to the beginning of this study (of which two were detected). This reflects the ongoing revolution that has been occurring since HTS became commonplace in plant virology and its subsequent inching into the regulatory sphere (Gauthier et al., 2022, Macdiarmid et al., 2013). Sampling and RNA extraction of tissue from all purchased plants, whether potentially symptomatic or healthy-appearing, allowed detection of viruses which may be latent or in the pre-symptomatic stage. A survey of peas in the UK using HTS found soybean dwarf virus in asymptomatic material but not symptomatic material, for example (Fowkes et al., 2021b). It is an important consideration to sample apparently healthy plants alongside potentially diseased ones. Where a plant showed what could be viral symptoms, this potentially symptomatic tissue was sampled, however, mostly so as not to 'miss' anything that could be causing symptoms. Each individual sample and all pooled samples contained apparently healthy tissue as well.

In the above survey of British pea leaves, re-sequencing pea samples at six-fold greater depth did not change the virus species composition found, though it did lead to recovering more of the genomes of the viruses that were detected (Fowkes et al., 2021b). In the datasets presented in the present chapter, this greater recovery of the genome only occurred when the reads from an individual plant extract sequenced at higher depth were combined with the reads derived from RNA extracted from the 27 pooled plants from the same location. Naturally this approach must be undertaken with caution so as not to increase the risk of assembling chimeric sequences not present in nature, but given that the pooled total RNA extract contained nucleic acid sequence from the same sample, it was deemed to be appropriate in this instance to combine the reads. Such an approach might be considered when a putative novel virus is detected in order to recover more of the genome, when higher sequencing depth alone is not sufficient.

However, given that RNA viruses have an error-prone replicase, they are likely to exist in the host as a swarm of related sequences within a host; a quasispecies (Domingo, 2002). This swarm vacillates around a master sequence which is most common, with one or a few dominant haplotypes and a smaller quantity of sequences which diverge from this main haplotype (García-Arenal et al. 2001). Viral genomes are present at very low amounts in a host sample; even with ribosomal RNA depletion, a PCR step is necessary to amplify viral sequence in the extracted nucleotides, which can introduce errors that are difficult to differentiate from 'real' SNPs, alongside the error rate of Illumina sequencing-by-synthesis itself, which is equivalent to viral replicases (Pérez-Losada et al., 2020).

Generally, HTS in plant virology is used to reconstruct a consensus or 'master' sequence from the highly variable population within a plant (Pérez-Losada et al., 2020), this is especially so for the sequences which emerge from pooled samples only, such as those representing the ONV1 genome. Flows of genomic molecules within an organism (i.e. systemic movement) over time can lead to population bottlenecks and turnovers that lead to evolution even within one individual plant, even without passaging into a new plant of the same or different species (Poirier and Vignuzzi, 2017). This bottleneck size can vary from a very small sub-population in some viruses such as potato virus Y to hundreds in the likes of cauliflower mosaic virus (Fabre et al., 2012, Monsion et al., 2008). All of these factors mean that the exemplary sequences provided here are something of an average of the virus sequences derived from the pooled or individual plant samples, or indeed the populations within individual plants, which differ amongst themselves. They are simply the *first* sequences found that belong to what are tentatively referred to as novel species and may not be representative of the true, theoretical population.

3.2.3. Cucumber mosaic virus

The dataset derived from FR-BLK (a pool of samples of French origin) contained a sequence 326bp in length which had 98.46% aa identity to the coat protein (CP) of cucumber mosaic virus (CMV). The presence of this virus was not confirmed in the plants extracted and pooled for this sample by either ELISA or qRT-PCR of the 9 plants sampled, which were negative for CMV. The healthy *N. tabacum* control did not return any CMV sequences. However, sequences were found matching other plant species sequenced on the machine at the same time, such as *Solanum* members.

Table 3.1 A table summarising the novel and known viruses detected in each sample sequenced via HTS. Read counts before and after trimming are given as number of paired reads. The finding of CMV in reads deriving from FR-BLK is likely to be contamination, see section 3.3.

Run	ArMV	UlPolV1	ONV1	ONV2	OVX	OAV1	OCV1	OACV1	OCV1	Oph	CMV
L15S20		✓			✓	✓	✓				
L15S6		✓			✓			✓			
L15S7		✓								✓	
L15S8		✓		✓	✓			✓			
VMS6					✓	✓	✓				
VMS18					✓	✓					
L17S14	✓	✓			✓	✓	✓	✓			
L17S17		✓									
L17S18		✓		✓		✓		✓		✓	
L17S19		✓			✓			✓			
L20S11					✓	✓		✓			
L20S12	✓										
L20S13		✓									
L21S1		✓									
L21S2		✓									
L21S4	✓										
L21S5						✓					
L22S3				✓							
L22S6				✓	✓		✓				
L23S3				✓							
IS-BLK				✓							
PL-BLK	✓	✓	✓			✓			✓	✓	
FR-BLK	✓				✓	✓			✓		✓
PL24			✓								

3.2.4. Ullucus polerovirus 1

Ullucus polerovirus 1 (UlPolV1) was previously detected in oca following HTS investigation of oca samples available via the internet in Europe (De Jonghe et al., 2022). sequences with species-level homology to UlPolV1 were found in the PL-BLK dataset and a number of datasets originating from nucleic acid extracts of other individual oca tubers (see Table 3.1). When confirmatory testing was undertaken on new extracts of total RNA from the plants forming PL-BLK using qRT-PCR, 9 plants out of the 27 that formed the PL-BLK sample tested positive. These plants were of an unknown variety and originated from an eBay purchase from a seller in Poland.

Further work was undertaken to characterise the UlPolV1 sequences derived from sequencing of RNA extracted from the oca samples. Two out of the four UlPolV1 nucleotide sequences on GenBank are partial, present as coat proteins only. A number of UlPolV1 sequences retrieved from the HTS datasets contained sequence matching to the CP region (ORF3). These were aligned using MUSCLE and the alignment is shown in Figure 3.3 on Page 58.

MUSCLE alignment was used to compare all four GenBank sequences with these two sequences. This revealed that some residues are conserved within but different between the groups of isolates from the two hosts, a number of which have different hydrophobicity. The oca-derived sequence shares approximately 85% nucleotide identity across its entire genome with the sequences on GenBank that were detected in ulluco, rising to 91% identity in the CP region. The predicted amino acid sequence of the P1P2 fusion (replicase) protein from the UlPolV1 isolate from oca is slightly more divergent, at 84% amino acid identity.

The UlPolV1 isolates from oca have a >10% amino acid difference in the P1-P2 fusion gene product from would indicate a new species according to ICTV demarcation criteria (Sõmera et al., 2021). However, all other gene products from the same virus are within this threshold demarcation threshold. For the time being the virus is presented here as an isolate of UlPolV1 because of the similarity in the coat protein gene product, the phylogeny (see subsection 3.2.4), and the predicted genome organisation. It is possibly, however, a member of a novel species.

UlPolV1 is relatively newly known to science, found via HTS of symptomatic *Ullucus tuberosus* plants in the UK in 2019 (Fox et al., 2019). Unpublished data showed little success transmitting this virus in ulluco plants using *Myzus persicae*, but its likely member of the genus *Polerovirus* would suggest that it UlPolV1 vectored by an aphid of some species and should be phloem limited (Hoffmann et al., 2001). At present, the symptoms of UlPolV1 and the threat of this virus to commonly grown crops in Europe is not well understood. UlPolV1 is serologically related to potato leafroll virus (PLRV), to the extent that ELISA antisera reacted to ulluco material that was deemed to be infected with UlPolV1 during an outbreak in the UK (Fox et al., 2019). It is thus possible that UlPolV1, which clusters based on host in oca and ulluco, is a South American virus which has come to Europe with the plants. Had the virus emerged from European flora and onto introduced oca and ulluco, one might have expected them to be more similar to one another, despite the difference in host, and the likely detection of this virus using serological methods in the Andes (see chapter 1) would point to this hypothesis.

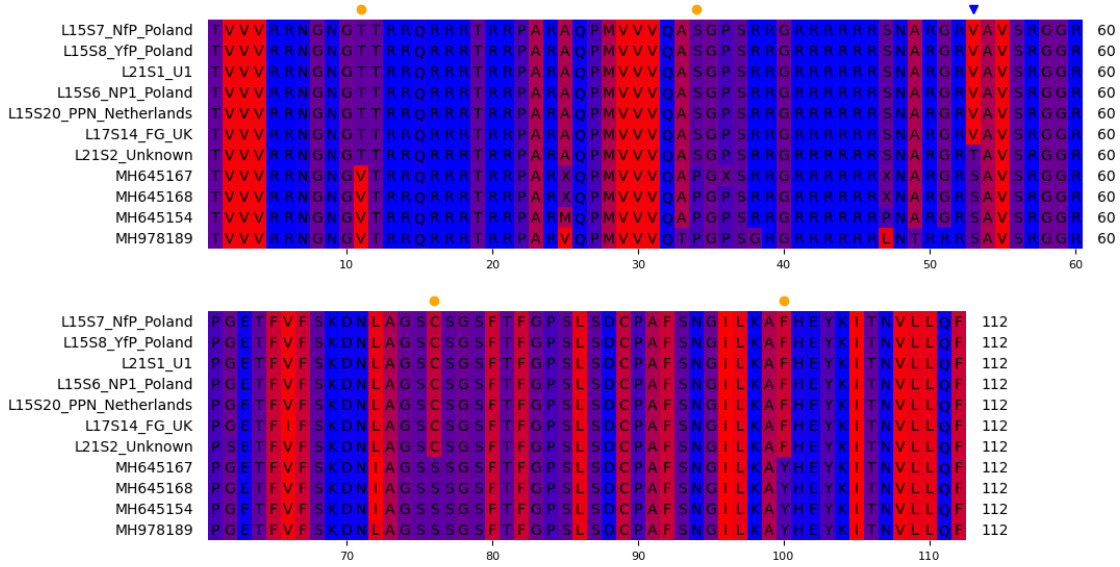


Figure 3.3 Alignment of the UlPolV1 ORF3 N-terminal regions coloured for hydrophobicity (red/paler is hydrophobic, blue/darker is hydrophilic). Sites that are consistent within sequences from the same host but differ between sequences from different hosts are marked with an orange circle. The residue marked with the blue arrow shows where the sequence detected in the HTS dataset L21S2 differs from all other sequences in this residue (threonine), but otherwise oca sequences code for a valine residue and ulluco sequences code for a serine at this position.

3.2.5. Arabis mosaic virus

A total of five HTS datasets (L17S14, L20S12, L21S4, PL-BLK and FR-BLK) contained Arabis mosaic virus (ArMV) sequences. When compared with existing ArMV sequences on GenBank using BLASTN, the ArMV sequences from oca shared 89-93% nucleotide identity with sequences of known ArMV isolates, suggesting a divergent isolate of ArMV in the oca plants. The GenBank sequences with the highest match (as measured by identity and coverage) to the oca isolates are accessions primarily infecting *Narcissus sp.* where the isolation host is indicated. Exceptions to this are sequences from FR-BLK (French and Polish origin) which shared 89% nt identity with accession GQ369526 from winter barley, and another with 91% identity to accession MH427275, isolated from *Apis mellifera*. An ArMV sequence from the PL-BLK HTS dataset (Polish and French origin) matched with an isolate from *Ligustrum vulgare* (accession EU617326) at 89% identical to one another to an isolate from *Fragaria sp.*, accession OR477267, with 94% nucleotide identity.

Where sample origin of these accession was indicated by associated metadata, the accessions with close homology to the ArMV sequences from oca are European. The exception is the *Apis mellifera* isolate above, accession MH427275, from Australia, which was sequenced as part of a study surveilling for plant viruses in bees (Roberts et al., 2018).

The plants that tested positive for ArMV using a qRT-PCR assay did not test positive when tested via ELISA using commercial antisera.

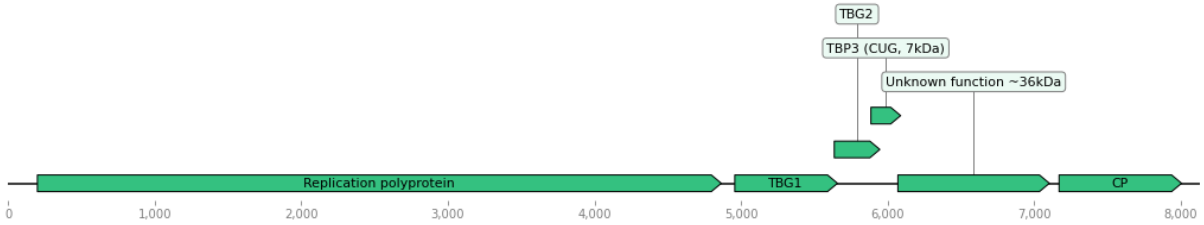


Figure 3.4 Putative genomic organisation of OAV1.

Large satellite RNA In addition to this finding of ArMV, one 1093bp sequence was retrieved from the FR-BLK dataset which matched with 89.7% nucleotide identity to the D0064 (exemplar) accession of Arabis mosaic virus large satellite RNA. This suggests that at least one plant whose RNA was extracted to form the FR-BLK pooled sample was infected by a likely divergent isolate of Arabis mosaic virus large satellite RNA.

3.2.6. Oca allexivirus 1

Sequences from nine HTS datasets (eight one-plant samples and the pooled FR-BLK) matched allexiviruses garlic virus X and garlic virus D with somewhat distant homology (72-74% nt identity), suggesting a novel species. These sequences clustered together and likely represent one species. What is likely to be the entire genome was recovered from several HTS datasets: FR-BLK, VMS18 and L20S11 (Pink Perfection, UK), L17S14 (unknown variety, free gift, UK), and L15S20 (Pink Perfection, Netherlands). All viruses detected in the HTS datasets are summarised in Table 3.1.

Primers were designed based on the consensus sequence of the virus as derived from the dataset FR-BLK, which was formed from a pool of samples originating from France. The genomic organisation presented here is based off the master sequence from the L20S11 dataset as it was the longest and is coding complete, with small 5' and 3' UTRs, and comes from a single plant rather than a master sequence of a pooled dataset. All sequences detected in oca plants which represent OAV1 share greater than 99% nucleotide identity with one another. All ORFs were in the same relative locations when the sequences were aligned using MUSCLE and did not differ between isolates. Variation between geographic isolates was no greater than between isolates from the same geographic origin.

qRT-PCR primers designed to the OAV1 sequence produced a positive result in total RNA extracted from freeze-dried tissue from one of the 2019 plants ('Pink Perfection', either HHPP or PPN).

The sequence retrieved from these datasets was predicted to contain six ORFs. Acarallexiviruses genomes code for a seventh ORF, which is a nucleic acid binding protein (NABP) ORF. This is absent in all known non-acarallexiviruses. The TGB3 gene of OAV1 alsohas no AUG start codon, instead beginning with an alternative CUG codon as in acarallexiviruses, blackberry virus E, and Arachis pintoi virus (Mansouri and Ryšánek, 2021). The predicted protein product of TGB3 is 7kDa, closer to acarallexiviruses in size than the other allexiviruses (Mansouri and

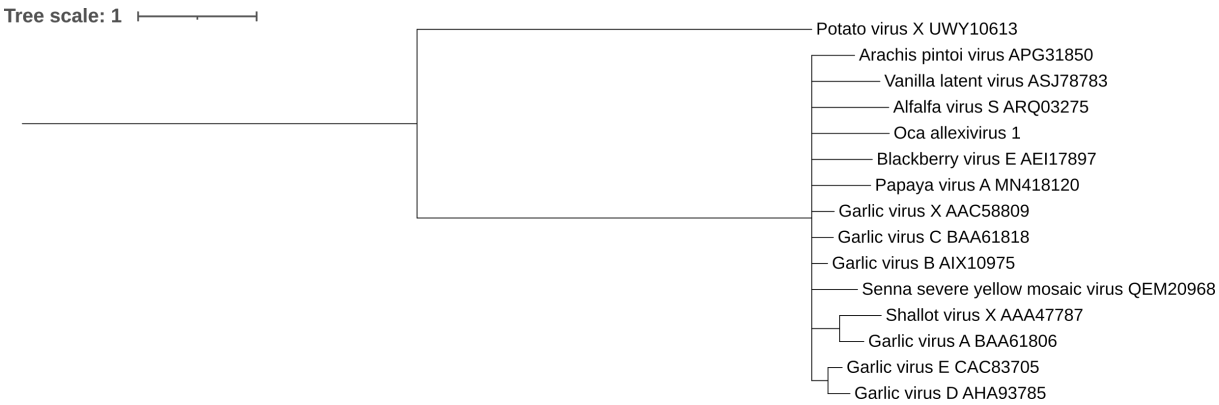


Figure 3.5 Phylogenetic tree based on the amino acid sequence of the replicase protein of OAV1 and other allexiviruses.

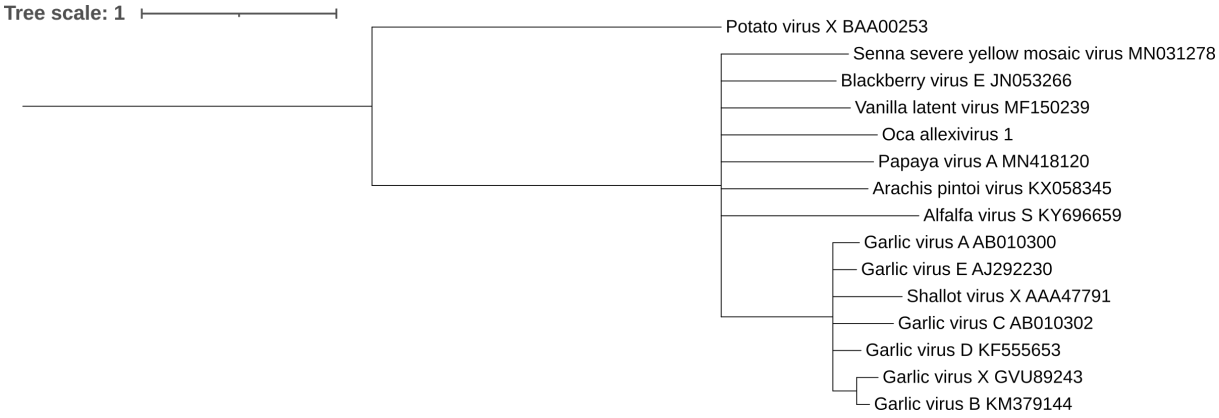


Figure 3.6 Phylogenetic tree based on the amino acid sequence of the predicted coat protein of OAV1 and other allexiviruses.

Ryšánek, 2021). The conserved protein of unknown function in ORF5 is predicted to be only 36kDa, closer to the unratified member papaya virus A than to other ratified allexiviruses, which typically have a 40kDa product from this gene (Read et al., 2020).

The demarcation threshold for allexiviruses is 80% amino acid identity or less in the CP or replicase genes (Kreuze et al., 2020b). Based on the phylogeny and genome organisation, OAV1 is likely to be a new member of a new species in the genus *Allexivirus*. The suggested species name is *Allexivirus oxalidis* (oca allexivirus 1, OAV1).

The phylogeny of OAV1 was inferred by aligning the predicted amino acid sequence of the CP and replicase genes with those of other allexiviruses. The resulting phylogenetic trees are shown in Figure 3.6 and Figure 3.5. As suggested by Mansouri and Ryšánek (Mansouri and Ryšánek, 2021) and later ratified, the genus *Allexivirus* splits into two sub-genera on this tree, and OAV1 sits within the non-acarallexivirus group, i.e. does not form a clade with the acarallexiviruses. Using either the CP amino acid sequence and the replicase polyprotein amino acid sequence, no strong clustering is supported within this sub-genus with the methods used, but the novel allexivirus does sit within the branch formed by other allexiviruses when placed on the tree using either gene.

The absence of a strong (i.e. supported by both bootstrap methods) acarallexivirus branch in the replicase tree (Figure 3.5) is likely due to the smaller number of sequences used to infer the

Table 3.2 A table of sequences retrieved from the oca HTS datasets and their amino acid identity to ophiovirus accessions on GenBank.

sequence	Len	RNA	Region	Accession	Virus matched	% aa
A	3768	RNA1	Partial RdRp	AWJ63454	Ranunculus white mottle ophiovirus	43.4
B	3496	RNA1	Partial RdRp	DBA06855	Adonis ophiovirus	24.3
L17S18C	2380	RNA1	Partial RdRp	DBA06952	Silene ophiovirus	41.5
L20S11H	1119	RNA1	Partial RdRp	DBA06952	Silene ophiovirus	52.4
L15S8D	1468	RNA3	CP complete	UUT43461	Citrus psorosis virus	26.9
L20S11E	1457	RNA3	CP complete	DBA06954	Silene ophiovirus	38.2
L15S7I	1532	RNA3	CP complete	DBA06954	Silene ophiovirus	38.5
L15S20L	1455	RNA3	CP complete	DBA06954	Silene ophiovirus	38.9
L17S18M	1441	RNA3	CP complete	UUT43461	Citrus psorosis virus	26.9
L15S7F	1602	RNA2	MP complete	DBA06953	Silene ophiovirus	39
L20S11J	1587	RNA2	MP complete	DBA06953	Silene ophiovirus	39.1
L17S18G	1669	RNA2	MP complete	DBA06953	Silene ophiovirus	39.6
L15S8K	1588	RNA2	MP complete	DBA06953	Silene ophiovirus	39.6
PL-BLK	233	RNA1	Partial RdRp	DBA06923	Lepidozia ophiovirus	60.8

*Primers match this.

** BLASTX showed higher coverage to some, less related ophioviruses, and lower coverage with higher identity to others.

phylogeny compared to Mansouri et al. who used several isolates per species to infer the tree (Mansouri and Ryšánek, 2021). The non-acarallexivirus branch is evident even with stringent boostrap parameters when using the coat protein sequence (Figure 3.6). Nonetheless in both trees, the novel virus clearly clusters with other allexiviruses.

The replicase of OAV1 is 61.3% identical to the replicase of senna severe mosaic virus. The coat protein is 43.7% similar to the coat protein of blackberry virus E.

A search on Serratus found no same-species hits, nor any SRA entries containing reads with identity greater than 90% to that of OAV1, but it did identity a catalytic RdRp motif.

The vegetative propagation of the plants infected with OAV1 is in line with other non-acarallexivirus members of the genus *Allexivirus*, which are known to spread chiefly via germplasm exchange with no other known vector at present. Acarallexiviruses are typically mite-vectored (Mansouri and Ryšánek, 2021). One plant infected with this virus arrived with mottling of the leaves. However, the plant had been sent via the post and was tuberising, so it is considerably more likely that the mottling observed was due to senescence or the abiotic stress of transport.

3.2.7. Ophioviruses

Table 3.2 on Page 61 shows a summary of sequences and the datasets they were obtained from. Two of the sequences presented are scaffolds of sequences which were later aligned using Geneious Prime. It is advisable to interpret scaffold A with some caution, as they come from datasets which originate from sequencing of RNA extracted from different samples (L17S18 and L15S8). However, since both are variety 'Yellow from Peru' samples from the same purchase

and only differ in two nucleotides, they have been included them as one scaffold for comparison. Scaffold B comes from two sequencing runs of the same plant, variety 'Pink Perfection' (L20S11, VMS6), and is thus almost certainly two sequences representing the same isolate. In no cases was the entire RdRp (RNA1) recovered.

Based on sequence information, there are at least two novel ophioviruses infecting specimens of *Oxalis tuberosa*. The tentative sequences given here are the result of assembling sequences from multiple different runs, as summarised in Table 3.2. Alignment of the sequences representing the CP gene suggest 64% similarity between the two isolates, well below the 85% species demarcation threshold suggested for by ICTV (García et al., 2017). No sequences representing RNA4 were found.

Only one HTS dataset contained sequences representing all three segments of one ophiovirus without a segment from another, closely related virus (either a member of the same species or another closely related species) also being present. That is to say, it was possible in this dataset alone to confidently associate three segments with one another as likely deriving from an infection with one virus. Thus it is highly likely that L17S18G/L15S8K (which are identical in amino acid sequence), L17S18M/L15S8D, and the A scaffold represent sequence of three segments of one virus. The virus is given here as oca ophiovirus 1 (OOV1). For the other virus(es), mixed infection makes it more difficult to associate one segment with another. The dataset from L15S7, which provided movement and coat protein coding sequences, did not produce any sequences matching to replicase coding sequence, though it suggests that sequences I and F belong to the same virus. As not all ophioviruses have a fourth segment (García et al., 2017), it is likely that none of the novel ophioviruses presented here have an RNA4, but it is not possible to say this for certain without further work.

In addition, the likely mixed infection of individual plants with different ophioviruses makes presenting a single genome difficult, both from a logistical point of view and a biological one. Logistically, without isolating the individual viruses from one another, there is no means by which the RNAs, which have no common motifs between one another, can be associated. Biologically, they may reassort anyway in nature (Thekke-Veetil et al., 2015). Phylogenetic analysis of the ophioviruses detected in oca suggested that they cluster strongly with one another in both MP and CP coding regions, making distinguishing the viruses through their relatedness impossible as well. For this reason, only one novel ophiovirus here is presented with a name, the other two MPs and CPs respectively are simply termed *Ophiovirus sp.* until further work can be undertaken to distinguish and study them. These sequences were all detected in the same ecological niche (oca plants) despite their genomic variance. This might be grounds for considering them as three variants of one species (Van Regenmortel et al., 2013).

The N terminal of the coat protein differs the most between the isolates, with the latter three quarters sharing only 80-82% amino acid identity, which is close to the species demarcation threshold set out by ICTV. Based on this information, at least one novel virus species can be proposed. The species name *Ophiovirus oxalidis* (oca ophiovirus 1, OOV1) is proposed for the

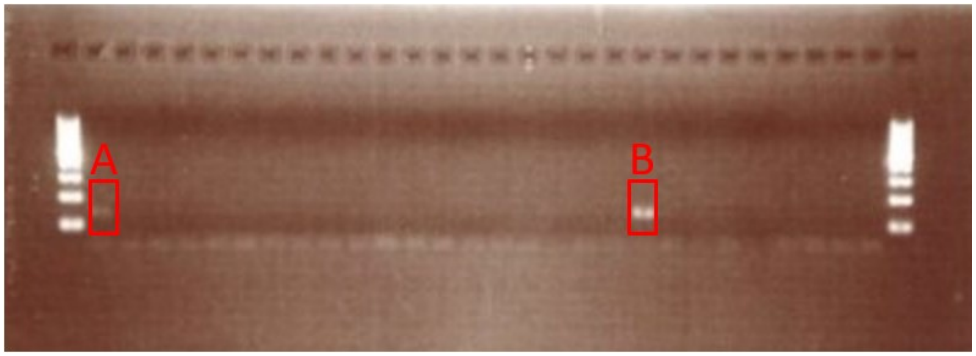


Figure 3.7 PCR products of ophiovirus-positive samples as tested by the generic ophiovirus primers from Vaira et al. 2003. A = weak band PL01 B = Strong band PL20.

three associated segments here. The other sequences may represent divergent isolates of the same species or sequence of another, closely related species in the genus *Ophiovirus*.
Generic ophiovirus primers (Vaira et al., 2003) successfully amplified RNA from four samples (Figure 3.7). Though qRT-PCR primers were designed to the original sequence, they were inconsistent in amplifying samples and as a result were not used in the follow-up survey.
As with the original finding paper, the node formed by the coat protein sequence of Buxbaumia ophiovirus has an unusually long branch length, greater than 9, which was flagged by IQ-TREE as unusual (Debat et al. 2023). For all three viruses, both the CP and MP proteins cluster a) with Silene ophiovirus and then b) with each other. This may indicate that they are divergent members of the same species, or simply closely related members of different species. The replicase gene was not compared with other ophioviruses as this gene was not recovered in its entirety from any oca isolate.
All of the ophiovirus sequences detected in oca are closely related phylogenetically. The closest known relative based on both CP and MP ORF products is Silene ophiovirus, from a paper mining for ophiovirus sequence in SRA archives (Debat et al., 2023). The SRA dataset that Silene ophiovirus was first described from derives from RNASeq of a specimen of *S. diclinis* from Spain (Muyle et al., 2021). No further biological information is available for this species, being present as mined sequence data only.
Ophioviruses are members of the family *Aspiviridae*. They are likely to be vectored by fungi, specifically a small number are vectored by filamentous *Olpidium* species members; the rest are unknown (García et al., 2017, Debat et al., 2023). No reads were identified as originating from members of the genus *Olpidium* in either dataset where a novel ophiovirus was detected.
However, both datasets contained assembled sequences which were assigned by MEGAN to the genus *Aspergillus*. The PL-BLK dataset contained sequences assigned by MEGAN to the species *A. parasiticus*, members of which can be plant pathogenic, such as in peanuts (Guo et al., 2008). Member of *A. parasiticus* can also act as contaminants of stored grains where they act as saprophytes (Okun et al., 2015). Further work would be required to determine if *A. parasiticus* is pathogenic within the oca plants and, if so, if it may be acting as a vector to ophioviruses.
Ophioviruses as a genus contain several members with economically important symptoms, such as big vein symptoms in lettuce (Roggero et al., 2000). A recent paper using SRA data increased

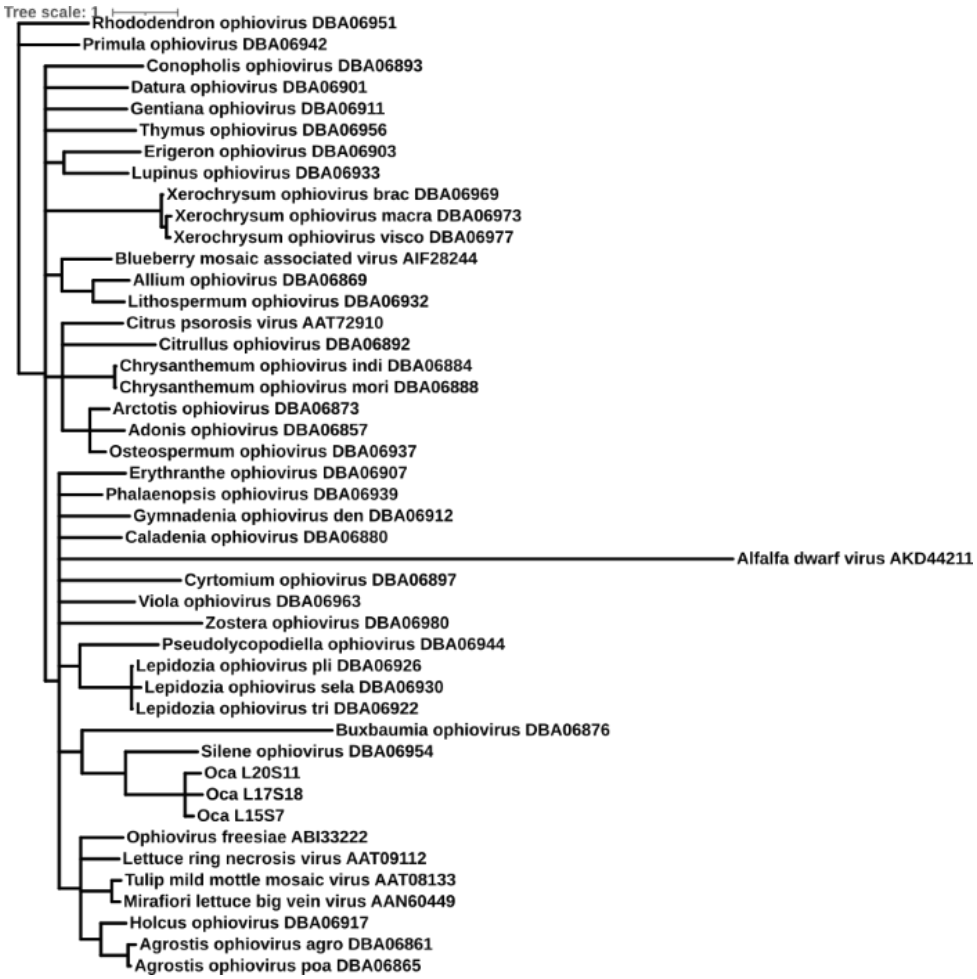


Figure 3.8 Phylogenetic tree placing the novel ophioviruses within the genus based on the coat protein amino acid sequence.

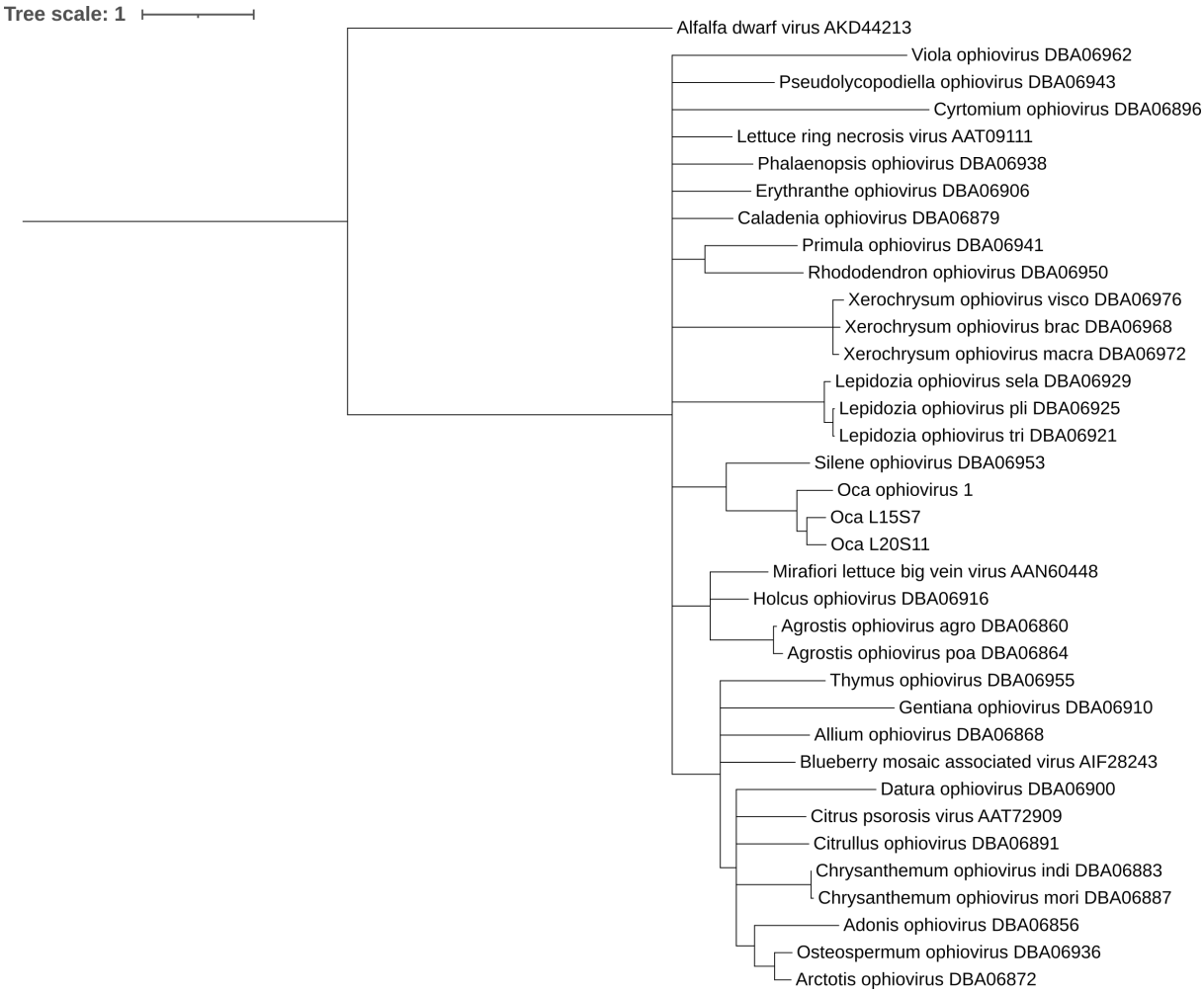


Figure 3.9 Phylogenetic tree placing the novel ophioviruses within the genus based on the movement protein amino acid sequence.

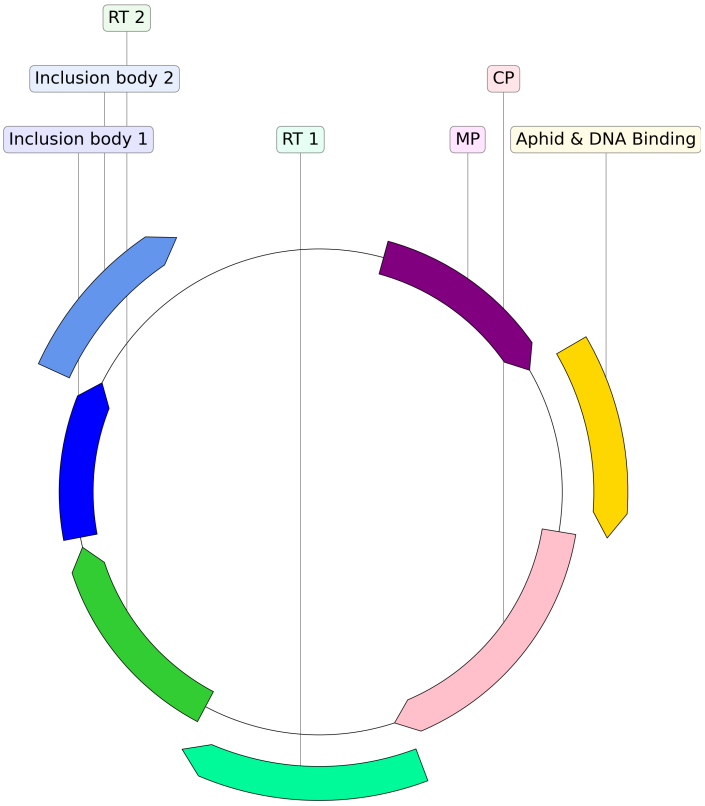


Figure 3.10 Genome organisation of the most complete sequence obtained for OCV1. This draft genome requires further confirmation as several open reading frames are atypical for caulimoviruses. The inclusion body (blue) and reverse transcriptase (green) ORFs are both broken into two by a stop codon. Similarly, the aphid transmission and DNA binding proteins are predicted to be fused into one ORF (yellow).

the number of known ophioviruses 4.5-fold, however, and many originated from asymptomatic samplings of wild hosts (Debat et al., 2023). Ophioviruses are also very diverse, some as little as 14% identical to one another in some regions of their genome (Debat et al., 2023). This could mean that sequences detected during this study and maybe others could be sequences of novel ophioviruses but are simply too divergent to detect using homology-based searches. For example, the Serratus database clusters species, genus and family level homology at 90%, 75% and 40% amino acid identity respectively (Edgar et al., 2022). Difficulty in detecting highly divergent viruses is a known constraint on homology-based searches such as BLAST (Kutnjak et al., 2021). The ophiovirus sequences from the HTS datasets derived from investigation of oca are highly divergent from those of other ophioviruses. However, sequences of CP genes from isolates of different ratified species may be only 31-52% similar to one another, with as little as 14.2% identity reported in one paper, so the finding of highly divergent ophiovirus sequences in oca would be congruent with this diversity (García et al., 2017, Debat et al., 2023).

3.2.8. Oca caulimovirus 1

Three sequences from the FR-BLK and PL-BK datasets showed distant homology members of the genus *Caulimovirus* (70% nt identity with 40% coverage). Some of these sequences were placed in the MEGAN 'root' node as the algorithm and BLAST detected that the sequences were similar to both cellular (plant) DNA and viral DNA accessions on GenBank. These three sequences varied in length from 4.6kb to 8.6kb. However, alignment of the three sequences indicated an assembly error that incorporated host sequence (or possibly host sequence attached to viral sequence endogenously). Removing the host-derived sequence information gave 97% identity between the sequences, suggesting the sequences belong to viruses that are members of the same species. The sequences were predicted to contain ORFs which coded for the movement protein, aphid transmission factor, DNA binding proteins capsid protein and replicase (reverse transcriptase polyprotein) products. No sequence in either dataset contained an ORF predicted to code for an inclusion body protein of appropriate size. Both sequences with this coding region present in the sequence suggested an appropriate start codon but the predicted ORF was split into two by a stop codon. This stop codon was in the same relative position in both sequences when aligned against one another.

Analysis showed that much of the genome of OCaV1 is represented by the recovered sequence and in mostly the canonical order for the family *Caulimoviridae*, but not entirely. The presence of reverse transcription, aphid transmission and translation activator sequences in these sequences does suggest that they may be replication-competent and even transmissible in theory, but without transmission demonstrated, this isn't certain. A phylogeny based on the putative replicase and coat proteins of a sequence from the FR-BLK dataset is presented here as the OCav1 genome.

Translation of the truncated ORF halves produced amino acid products which had distant homology to the translation activator protein of soybean Putnam virus (31.5% identity) for the 5' region, and for the 3' region matched plant hypothetical proteins and CaMV promoters (30% identity). All predicted ORF products aligned with other members of the genus *Caulimovirus* and produced relevant hits when checked for homology using BLASTP, though the reverse transcriptase of the caulimovirus from oca is only 65% similar to the reverse transcriptase of cauliflower mosaic virus (CaMV), which it is most similar to out of all caulimoviruses. The putative coat protein of OCaV1 is 39.2% similar to that of CaMV.

The draft genome derived from these datasets suggests that the aphid transmission factor of OCaV1 is fused to the DNA binding protein and is *after* the putative movement protein, which is not typical for caulimoviruses (Teychene et al., 2020).

DNA viruses which infect plants, such as members of *Caulimoviridae* (dsDNA) and *Geminiviridae* (ssDNA), are known to integrate into the plant genome as endogenous viral elements (EVEs) (Filloux et al., 2015b). The first integrated viral element found in plants was a geminivirus corresponding to its origin of replication and the replicase gene, where the mechanism is believed to be illegitimate recombination (Bejarano et al., 1996).

This process can lead to difficulties with homology-based methods of virus identification, as sequences may have high identity in small portions to 'plant' sequence which derives from a viral source somewhere in the past. This integration cause difficulties in differentiating between episomal (replicating) and integrated forms without further experimentation (Massart et al., 2019). In addition, because NCBI GenBank is not curated like some databases, sometimes individual accessions are incorrectly labelled as the wrong organism (Kutnjak et al., 2021). MEGAN, which uses the Last Common Ancestor (LCA) algorithm to assign reads to taxa (Huson et al., 2016), may be unable to distinguish sequences of this kind. MEGAN may be unable to determine whether sequences are plant (cellular) or viral in nature, and thus places them in the root, which is the earliest possible point in the tree as viruses and cellular life may not have a common ancestor at all. It remains the work of an individual investigator to determine if sequences assigned this way (or indeed those assigned to certain DNA virus taxa) represent virus or merely integrated sequence, potentially inherited from ancestral infections in the plant's lineage.

The difference between integrated and episomal forms of viruses is not always clear-cut from a risk management perspective. Integrated forms of DNA viruses can activate and begin replicating. In banana (*Musa sp.*), recently integrated badnavirus sequence, even when rearranged and fragmented, can become infectious upon hybridisation of the hosts, or wounding, and such rearranged sequences taken together can represent the entire genome of the badnavirus (Gayral et al., 2008). Investigation of yam (*Dioscorea alata*) through extraction of viral RNA using VANA with DNAse digestion followed by HTS led to the finding and investigation of transcriptionally active geminiviral EVEs; sequences similar to these were found in many *Dioscorea sp.* samples (Filloux et al. 2015b). Even if a person is reasonably certain that the sequence found originates from the host genome, presence of viral replicase sequence could then lead to an replicative form in future in some rare cases.

Dahlia mosaic virus strain DMV-D10, now known as Dahlia variabilis endogenous plant pararetroviral sequence (DvEPRS) is one example of an endogenous sequence (Eid and Pappu, 2014). In this way, some species of virus are known only through integrated sequences in plant hosts. DvEPRS sequences were found to be common in yacón HTS datasets in this thesis (see chapter 6 and subsection 4.2.3).

Neither phylogenetic tree inferred during the present study suggested a strong clustering of OCaV1 with any particular caulimovirus. However the OCaV1 genes branch with them as a genus, forming part of a broader cluster that contains all caulimoviruses except strawberry vein banding virus and Metaplexis yellow mottle associated virus.

The demarcation threshold for caulimoviruses is a difference of 20% or greater in the amino acid of the replicase (Teycheney et al., 2020). Regardless of whether this sequence is endogenous or not, it likely represents a new species of caulimovirus. The name *Caulimovirus oxalidis* (oca caulimovirus 1, OCaV1) is proposed. Further work is required to refine the draft genome and determine if the unusual organisation is genuine or an assembly error.

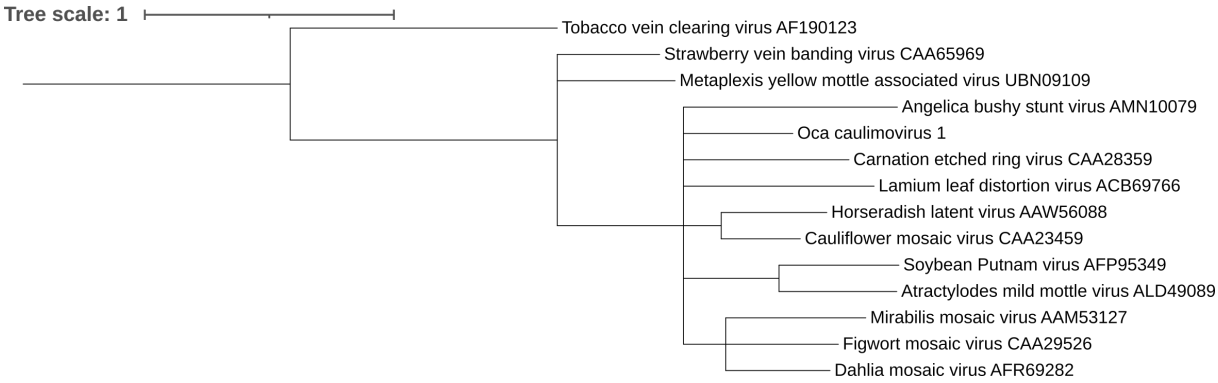


Figure 3.11 Phylogenetic tree of the oca caulimovirus 1 coat protein amino acid sequence compared with other members of the genus *Caulimovirus*.

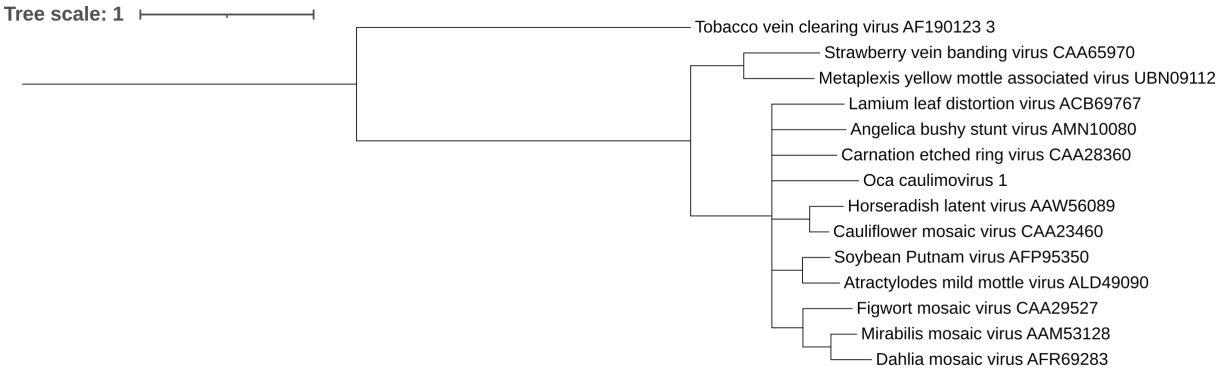


Figure 3.12 Phylogenetic tree of the oca caulimovirus 1 replicase amino acid sequence compared with other members of the genus *Caulimovirus*.

3.2.9. Oca capulavirus 1

Sequences from several HTS datasets showed distant identity (74.7% identity with 37% query coverage) to Euphorbia caput-medusae latent virus accession NC_075080, a capulavirus. The longest of these and likely to represent the entire genome was a sequence from the L15S20 dataset, Netherlands internet purchase, 2792 nucleotides long. This sequence (along with several others) contains the 'TAATATTAC' origin of replication motif common to all capulaviruses (Varsani et al., 2017). A self-alignment using BLASTN indicated that the first and last 106 nucleotides were identical, a common error in *de novo* assembly of circular viruses. Removing the first 106 nucleotides led to a 2686bp circular genome with the origin of replication and capulavirus ORFs intact. This virus is putatively named oca capulavirus 1 (OCV1).
This created the following ORFs: V1 (CP), V3 and V4 (movement proteins), C1, C2 (Replication proteins) and C3 (Bernardo et al., 2013). Apart from C3, which is contained entirely within C1 and thus matches C1 by homology searches, these were confirmed using BLASTP. V2 is absent from the OCV1 sequence and mapping of sequences to the putative genome using Geneious Prime did not suggest the presence of the ORF in any relevant sequence. The putative C1:C2 readthrough ORF does not produce a product divisible by three (i.e. translatable via codons), making it dubious, so it has not been included.
The ORF corresponding to the RepA (replication associated) gene has a 99.5% match at 76.81% query cover with accession JABWDY010003008.1, which originates from HTS sequencing and

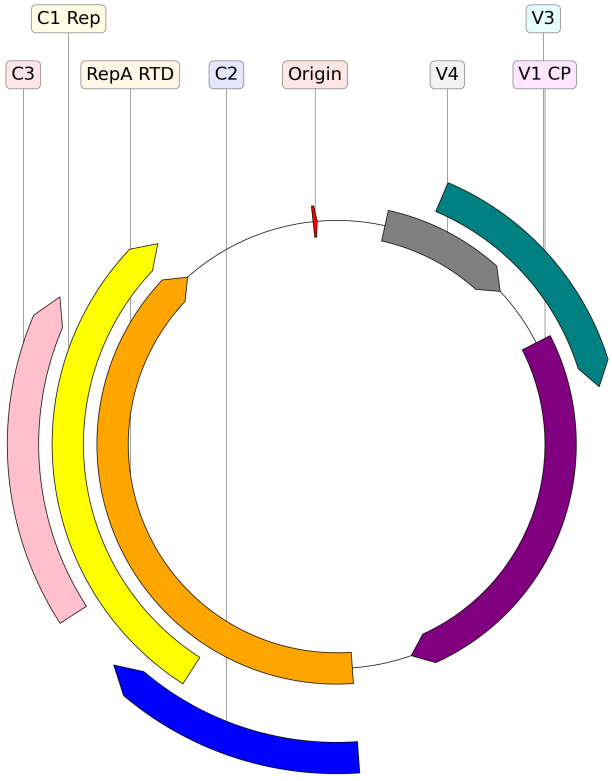


Figure 3.13 Putative genome organisation of OCV1, showing ORFs and the conserved replication sequence of geminivirids.

assembly of an extract from *Thalictrum thalictroides*. This sequence is listed on GenBank as the organism *T. thalictroides*, but its annotations suggest a geminivirus. Accession JABWDY010003008.1 is a 3125bp long sequence which aligns to the OCV1 draft genome with a 96% identity.

Alignment of the end of the capulavirus sequence represented by accession JABWDY010003008.1 (approximately 350bp) with the OCV1 sequence from oca using BLASTN suggests that the 5' end of accession JABWDY010003008.1 matches the 3' end, as with the sequence representing OCV1 sequence that was obtained from the oca datasets.

An *in silico* primer test using Geneious Prime of the OCV1 primers against accession JABWDY010003008.1 produced a hit. Additionally, while the JABWDY010003008.1 accession contains an insert in the putative C2 gene that matches no accessions on NCBI, it has the same open reading frames, including an absent V2 ORF before the V4 ORF. It is possible that the V2 gene is not functional in capulaviruses (Bernardo et al., 2016). It is thus very likely that *T. thalictroides* is also a host of OCV1.

qRT-PCR primers designed to this virus produced a positive result when used to amplify total RNA extracted from freeze-dried tissue of one of the plants sequenced via HTS in 2019 ('Pink Perfection', unknown origin).

Amplification of OCV1-positive samples using RCA and generic hexamers led to an amplification product with considerably lower Ct values when amplified by qRT-PCR than the original extract which was used for RCA (20.3/20.4 for the non-RCA extract control vs. 8.1/9.56 for the RCA product), suggesting successful amplification. The lower Ct values obtained when testing the RCA product compared with testing the RNA extracted from the plants does suggest that circular sequence may be present in the host cell extract, which is likely to be episomal. Likewise, the recovery of a full genome with intact ORFs from the HTS datasets from RNA extracted from different samples does suggest that the OCV1 sequence detected in the extracts from oca plants is not integrated, but without demonstration of transmission, this cannot be confirmed. Moreover, the primers were designed to the genome but may be detecting integrated sequence in plants.

The capulavirus sequence detected in the oca plants likely belongs to a species whose members were also sequenced in along with the transcriptome of *Thalictrum thalictroides*. Though NCBI lists the data as unpublished, they were later published as a paper on the transcriptomes of ranunculids (Arias et al., 2021) which links the SRA entry to samples of plants from the United States, specifically two accessions cultivated from nursery stock in Michigan and North Carolina. One of the plants listed in this BioSample is the vouchered specimen V. Di Stilio 124 (WTU), which according to another paper originates with a nursery in Michigan (Morales-Briones et al., 2019), who offer *Oxalis* species for sale as of 18/1/24.

The virus detected through HTS could have infected the oca plants sampled (or their progenitors) in Europe, given the presence of the virus in naturalised oxalids and ornamentals. Oxalis 'Irish Mist', which tested positive for OCV1, is listed by sellers and gardening websites as belonging to different species, some of which are South American. The sample purchased for

this study was listed as *O. triangularis papillonacea*. *O. regenelii*, listed on one US website (https://www.quackingrassnursery.com/plant/Oxalis-regnellii-Irish-Mist) is a synonym of *O. triangularis* (Govaerts et al., 2021) which is a South American species (Gardner et al., 2012), as is *O. articulata* as listed on the RHS website (The Royal Horticultural Society, 2024, Gardner et al., 2012). *O. articulata* is a naturalised non-native in Great Britain (Henning and Raab-Straube, 2016). One listed name for the variety 'Irish Mist' is South African (*O. depressa* (The Royal Horticultural Society, 2024, Oberlander et al., 2011)). Neither name listed by the RHS is the same as those listed by the two sellers mentioned here, so it is possible that the sellers have mislabelled the plants due to their large, triangular leaves.

Neither *O. depressa* nor *O. triangularis* were listed on the Euro+Med database as naturalised when queried. If the virus found was reconstructed from an EVE in the ancient ancestor of *Oxalis sp.*, even the South American branch (on which *O. tuberosa* and *O. triangularis* are quite far apart (Gardner et al., 2012)), then one would expect it to differ more substantially. While EVEs evolve much more slowly than 'live' viruses, they still evolve at a roughly equivalent rate to their eukaryotic hosts (Aiewsakun and Katzourakis, 2015). Capulaviruses have been found to have wide distributions before. Alfalfa leaf curl has been found in natural infections in both Europe and in Argentina (Hily et al., 2021). Accession JABWDY010003008.1 and the paper associated with the HTS experiment which produced this sequence suggests also that OCV1 is able to infect *T. thalictroides* and may be circulating in the United States. A population analysis of full genomes of OCV1 (or integrants) would help to eludicate the origin of the virus, especially as no RHS accessions from South America or South Africa tested positive for this virus.

Capulaviruses are thought to be inefficiently vectored by aphids, and thus have low incidence (Ryckebusch et al., 2022), though in this survey, one group of plants from the same origin was entirely infected by OCV1. This could suggest a slightly more efficient transmission process than previously assumed. Other viruses detected in these oca samples were not present at the same 100% incidence, i.e. they are presumably not clones of one another or tubers from the same plant.

Given this, it seems likely that OCV1 detected infecting these oca samples is a) actively replicating and infectious but b) not of concern. No symptoms were associated with the OCV1 infection.

This genome organisation of OCV1 appears to be different from established geminivirids, both those in genera and those not assigned a genus (Varsani et al., 2017). For this reason sequences of geminiviruses other than capulaviruses were also included on the phylogenetic tree constructed for OCV1. OCV1 is most similar in both RepA and CP genes with the RepA and CP genes of the other four ratified and one proposed capulavirus used for the tree, despite the atypical genome organisation, separate from those of mastreviruses and grabloviruses.

The replicase of OCV1 is most similar to that of Euphorbia caput-medusae latent virus at 68.8% identity. The coat protein is most similar to that of French bean severe leaf curl virus, at 58.2% identity. Across the genome, OCV1 is most similar in sequence to Euphorbia caput-medusae

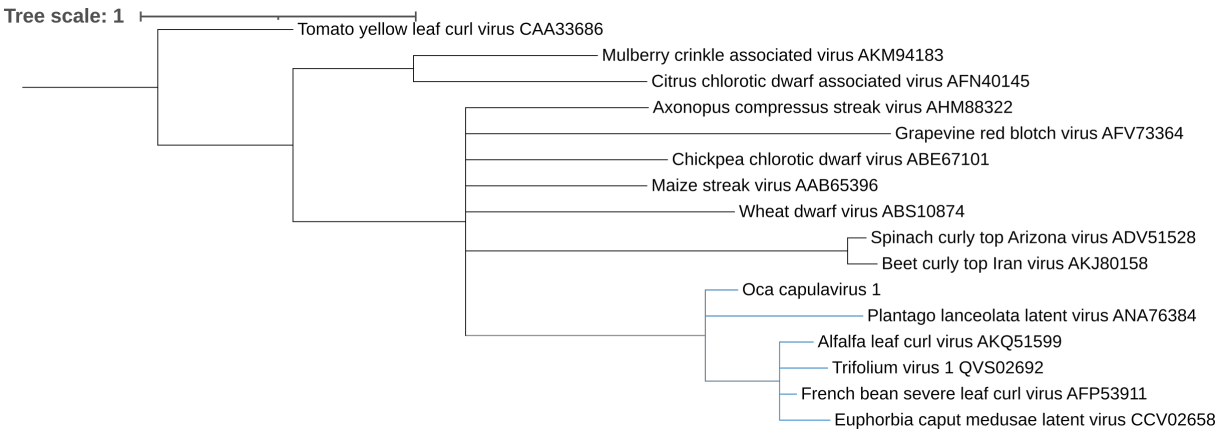


Figure 3.14 Phylogenetic tree of the oca capulavirus 1 coat protein amino acid sequence compared with other members of the family *Geminiviridae*. Capulaviruses are marked in blue.

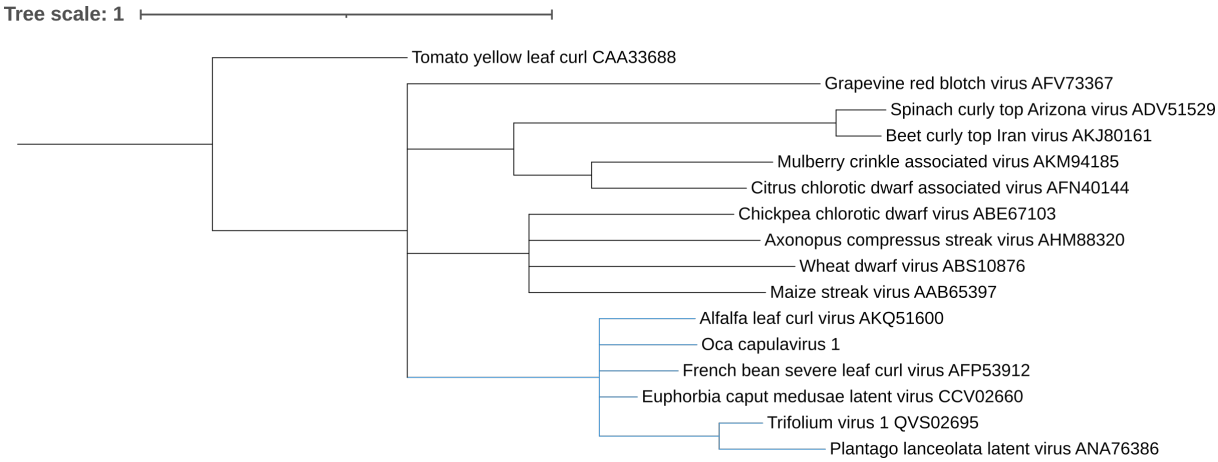


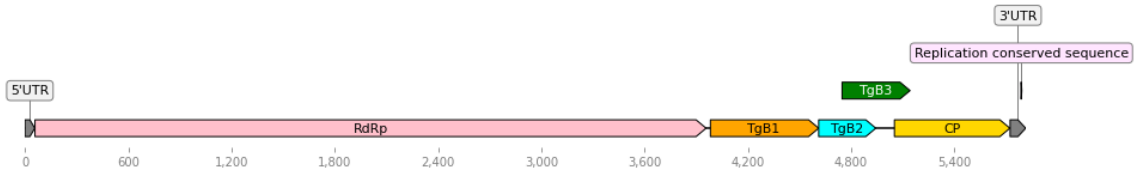
Figure 3.15 Phylogenetic tree of the OCV1 replicase amino acid sequence compared with other members of the family *Geminiviridae*. Capulaviruses are marked in blue.

latent virus, at 51.6%. The current ICTV species demarcation threshold for capulaviruses is 78% nucleotide identity (Varsani et al., 2017). Based on this, the name *Capulavirus oxalidis* (oca capulavirus 1, OCV1) is proposed for the species.

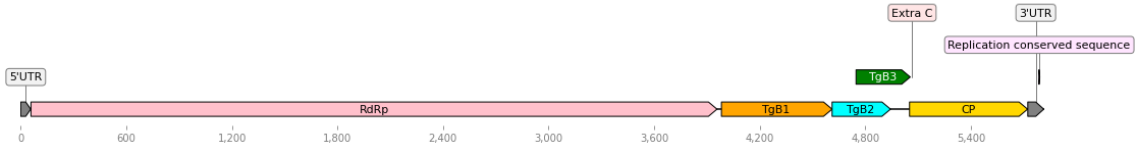
3.2.10. Oca virus X

Sequence matching with 72.8% nucleotide identity to yam virus X with very low (7%) coverage was recovered from extracts from the oca plants. Investigation of this sequence suggested a novel potexvirus, oca virus X (OVX). Analysis of the open reading frames suggested a relationship not to yam virus X, but to mint virus X. Oca virus X is most closely related to mint virus X (MVX) in both coat protein and replicase gene products. In terms of identity, the coat protein is 50.6% identical to MVX and the replicase is 60.5% identical in amino acid sequence to both MVX and cardamom virus X.

The putative genome of OVX contains the conserved potexvirus replication sequence on its 3' end, 'ACUUAA' (AC[C/U]UA) which according to RNAFold would be exposed on the end of a hairpin loop as in bamboo virus X (Cheng and Tsai, 1999) 3.17. The sequence representing



(a) Organisation of the OVX genome determined through HTS alone, using a sequence from the FR-BLK pooled sample HTS dataset as an exemplar. The only difference between this sequence and the consensus of the 99% similar sequences is the 'AGAAAA' sequence at the beginning of the 5' UTR. Note the longer TGBP3 putative ORF.



(b) Organisation of the OVX genome indicated by the same sequence with an additional 'C' added after Sanger sequencing, leading to a smaller TGB3 product more in line with other potexviruses.

Figure 3.16 Comparison of the genome organisation of OVX before and after correcting for the atypical TGB3 gene size.

OVX also contains the FDFFDG motif in the predicted coat protein sequence, common to some filamentous viruses (Dolja et al., 1991).

This sequence lacks the 'GAAAA' sequence conserved in almost all potexviruses (Chen et al., 2005), or any 'AAAA' 5' sequence, so it is likely not complete. However, it is likely to be *coding* complete. Forming a consensus from all available OVX sequences matching the French isolate did produce a slightly longer sequence starting with 'AGAAAA', but due to the mixed origins of the samples, such a consensus must be interpreted with caution, especially given the evident variation between the samples noted above. For this reason, the coding complete genome recovered from the FR-BLK dataset is presented here as the most complete genome.

The conserved potexvirus 'GUUAAGUU' promoter regions upstream of the CP and TGB blocks are both present in this sequence (Kim and Hemenway, 1997).

An unusual property of these sequences, shared with all of the OVX sequences derived from the HTS datasets, is the large TGB3-coding ORF4, which overlaps with the CP (frame +1) and is predicted to produce a protein between 13k and 14k kDa in weight. Sanger sequencing of amplicons obtained with primers designed to this portion of the genome produced this missing nucleotide in the forward sequence but not the reverse sequence. Editing the sequence to include this cysteine from the forward sequence resolved ORF4 to end just before the CP, producing a predicted 10kDa product. This product is much closer in size to the norm for the genus (Kreuze et al. 2020). It is possible that the primers designed to this region selectively amplified different sub-variants of the population within the host, leading to the difference between forward and reverse products.

A search on Serratus found no same-species hits, nor any SRA entries containing reads with identity to OVX sequence greater than 90%, but it did identify an RdRp catalytic core motif.

The qRT-PCR primers are likely to produce positive results the isolate detected in the pooled FR-BLK dataset according to *in silico* primer matching and detection in the plants, but likely not

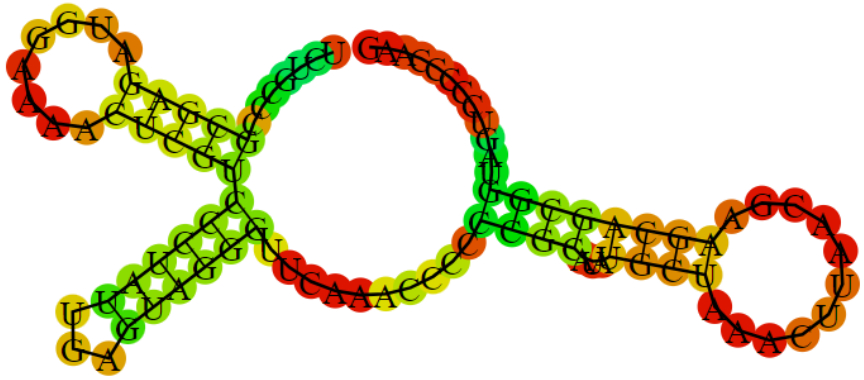


Figure 3.17 The predicted structure of the 5' UTR of OVX from RNAFold, showing a hairpin loop common to the genus.

the isolates detected in the 'Yellow from Peru' (L15S8) or 'Negra' (L15S6) datasets, at least according to an *in silico* analysis allowing for 1 mismatch.
The cnPCR primers are likewise designed to the FR-BLK-derived biotype and not to the other biotype. An *in silico* analysis suggested that they do not match every sequence generated from HTS of oca, but the primers exist for sequencing, not for detection. Indeed 2/3 of the samples whose RNA extracts were amplified for the generic potexvirus primers did not amplify with these species-specific primers.
The generic cnPCR primers from (Van Der Vlugt and Berendsen, 2002) produced a weak band at 280bp for DK2, and strong bands at 280bp for DK1 and B1 (Figure 3.18).
Phylogenetically, OVX clusters with mint virus X most closely, followed by a small clade containing cardamom virus X, Phaius virus X and lily virus X. This clade has a broad distribution of first findings. Cardamom virus X was first detected in small cardamom transcriptomic datasets from India (Sidharthan et al., 2021b), with no later reports suggesting a broader distribution at present. Mint virus X was detected in symptomatic mint and found to be widespread in nurseries around the US (Tzanetakis et al., 2006). Lily virus X infections have been reported in Europe, Asia, Argentina and Australia (CABI, 2024).
The sequences from HTS datasets L15S8 (Yellow from Peru, Poland) and L15S6 (Oca Negra, Poland) are divergent from the other OVX sequences (which share 99.6% nucleotide identity with one another) and clustered separately from the rest. These sequences did not align perfectly with each other. All sequences shared 88.4% nucleotide identity with one another overall. Alignment of the predicted replicase amino acid sequence showed that the sequences derived from RNA extracted from all five plants shared 97.7% pairwise identity, suggesting, as one might expect, selective pressure keeping the replicase product region similar with synonymous codons between the isolates which differed in nucleotide sequence. All sequences where the CP gene was predicted to be present (thus excluding L17S14) shared greater than 93.4% pairwise

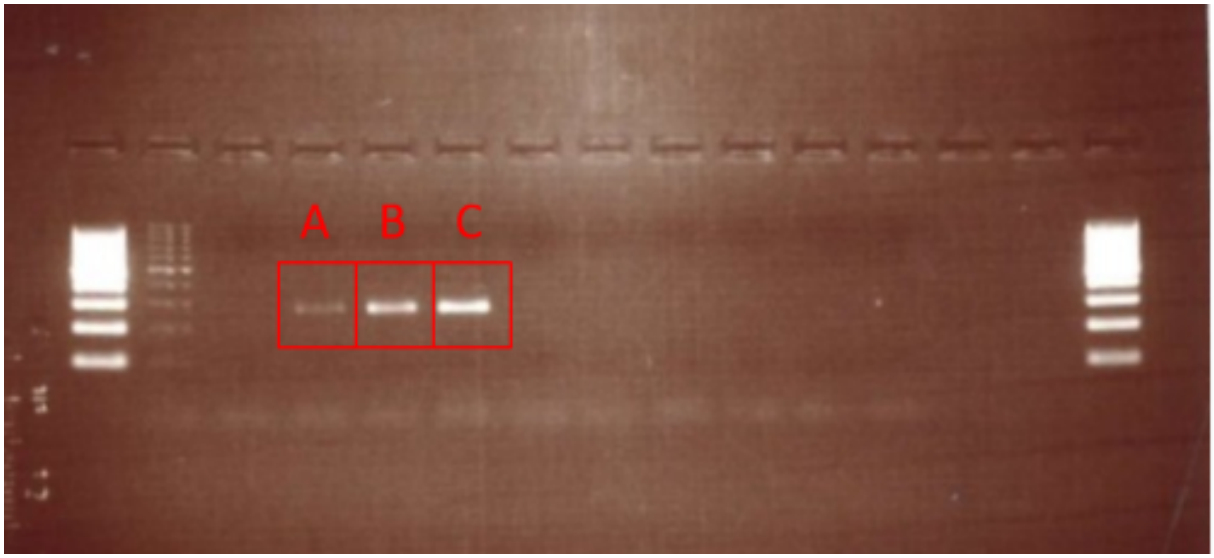


Figure 3.18 PCR products of OVX-positive samples when tested using generic potexvirus primers from Van Der Vlugt and Berendsen, 2002. A = DK2, B = DK1, C = B1.

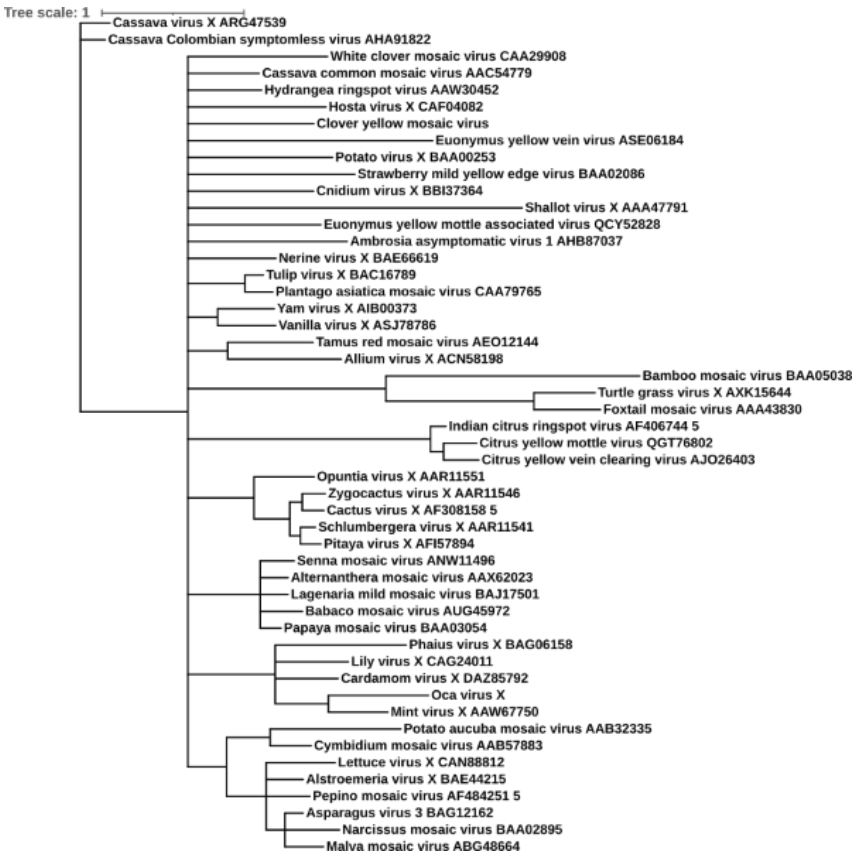


Figure 3.19 Phylogenetic tree placing the novel potexvirus within the genus based on the coat protein amino acid sequence.

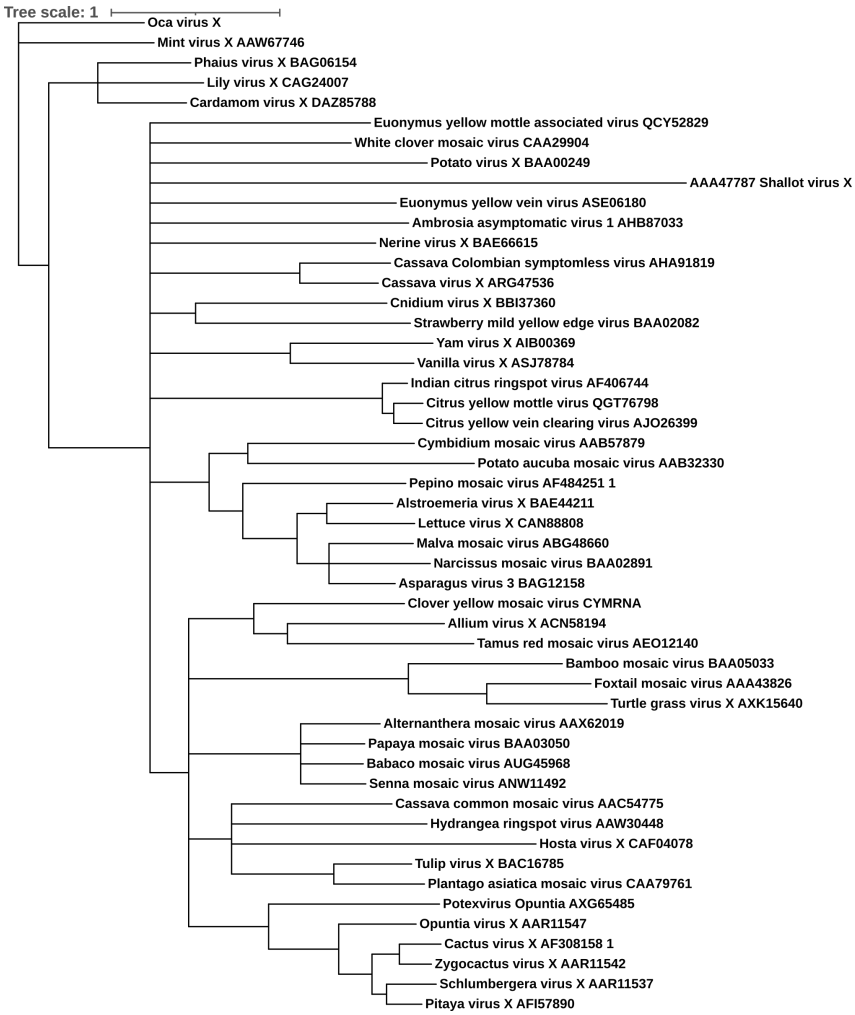


Figure 3.20 Phylogenetic tree placing the novel potexvirus within the genus based on the replicase amino acid sequence.

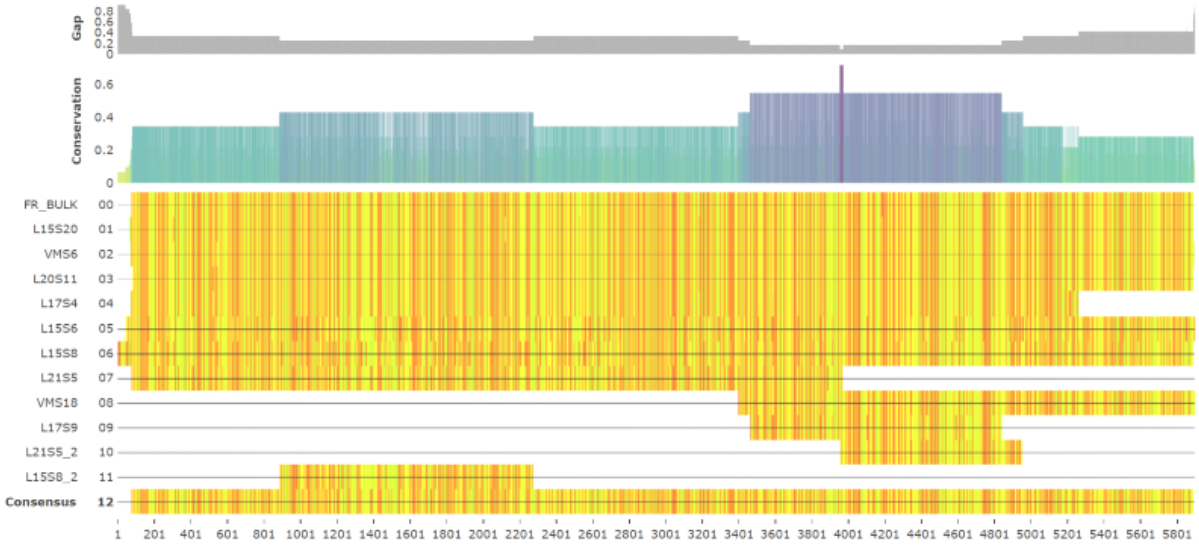


Figure 3.21 An alignment of the potexvirus sequences from from different datasets, coloured by identity. The isolates from the L15S6 and L15S8 HTS datasets are most different from the others, at 93% identity.

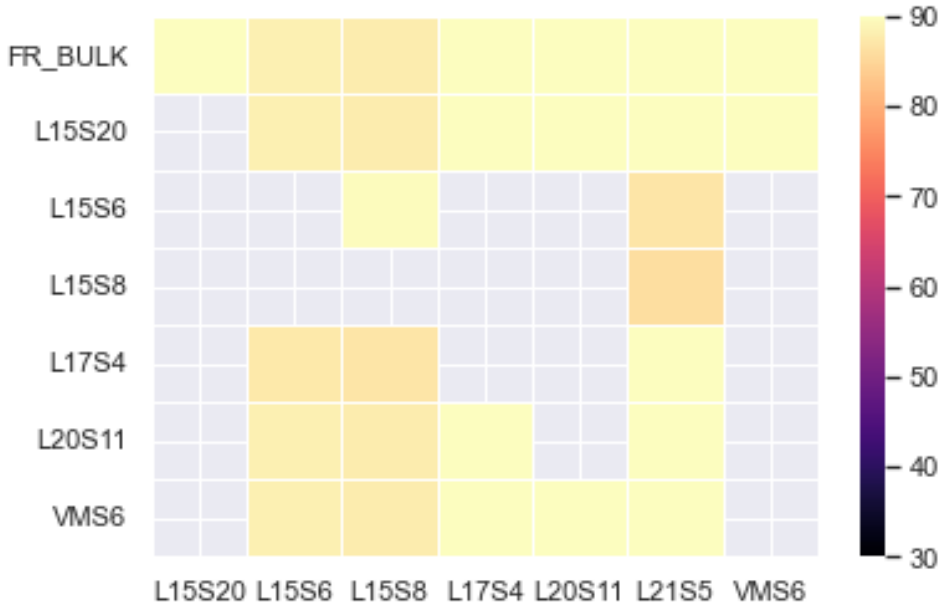


Figure 3.22 SDT output for all sequences labelled by the HTS dataset they originate from, excluding those present as only small fragments. Though all have high identity, VMS6, L15S6, L15S8 and L20S11 sequences differ from the others.

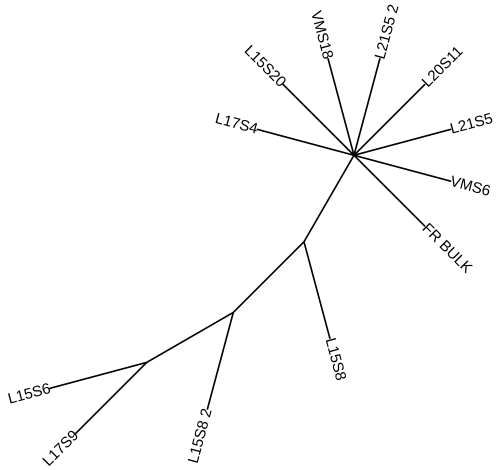


Figure 3.23 An unrooted phylogenetic tree based on the nucleotide sequences of OVX, showing that the OVX sequences fall into two main groups. This tree has been collapsed to only branches with high bootstrap support and so the most common biotype is present as one cluster.

identity and 97.8% amino acid identity with one another in this region. When compared across the nucleotide sequence of the genome, isolates L15S6, L17S9 and L15S8 clustered together, away from the other isolates (Figure 3.23). This suggests that there are two biogroups of OVX infecting these various oca samples. The VMS6/L20S11 (HHPP) isolate likely belongs to the more common biogroup despite some differences in sequence.
The ICTV species demarcation threshold for potexviruses is 80% amino acid identity or less in CP or replicase gene products (Kreuze et al., 2020b). Based on the genomic information and phylogeny, the virus here is suggested to be a novel member of the genus *Potexvirus*, suggested species name *Potexvirus oxalidis* (oca virus X, OVX).

3.2.11. Nepoviruses

Three different nepoviruses were detected in the oca plants sampled, two of which are likely to belong to novel species. These viruses are not the first nepoviruses found to infect oca. This is the second finding of ArMV infecting oca from Europe (De Jonghe et al., 2022). Potato black ringspot virus (PBRSV) has also been found to naturally infect oca (Jones and Kenten, 1981), though none of the plants tested during this study were found to be infected with PBRSV and infections of European oca with PBRSV have not been reported in the literature. Testing during this study detected UlPolV1 and beet ringspot virus (BRSV) in oca (subsection 4.2.2), new to this host, and Arabis mosaic virus. In general, members of the family *Oxalidaceae* were not commonly tested for any nepoviruses in the literature, except *Oxalis repens* which showed no symptoms when inoculated with TRSV (see chapter 5).
A number of grapevine-infecting viruses, including ArMV, CLRV and GBLV, are thought to have a Eurasian origin (Digiaro et al., 2017). BLMoV is thought to originate from the Americas and, like CLRV, is pollen transmitted with no known nematode vector (Rowhani et al., 2017). Stenotaphrum nepovirus, which is closely related to ONV1, was first detected in Australia (Tran et al., 2021). None of these viruses have been detected in oca from South America, though some nepoviruses were first discovered there, such as potato virus U (Jones et al., 1983). Further extraction and sequencing of viruses from a greater number of oca plants could detect further isolates of *Nepovirus primoxalidis* and help to elucidate the species' origin via phylogeny. This is complicated by the uncertain progeny of some oca plants available via the internet, however.

Arabis mosaic virus It isn't possible to say with any certainty from where the ArMV population detected in oca originates, though the species itself is likely to be European in origin (Digiaro et al., 2017). Comparison based on nucleotide identity is not sufficient and a phylogenetic tree should be generated (Duffy and Seah, 2010). Without ArMV sequences on GenBank that originate from Peru, Bolivia and Colombia, it cannot be said for certain that the relatedness of ArMV isolates from oca to European *Narcissus sp.* ArMV isolates reflects the most recent ancestry, or only that these are the most closely related isolates available as sequences on GenBank.

Table 3.3 Comparison of read depths and lengths before and after concatenating the file containing the pooled (PL-BLK) dataset reads with the reads from another, higher-depth dataset from sequencing of RNA extracted from a single plant originally pooled to and sequenced as part of the PL-BLK dataset.

(a) ONV1 sequences obtained from assembly of reads from the PL-BLK pooled HTS dataset.

Region	Depth (average)	Reads used	Length
RNA1	18.9x	503	5762
RNA2 3' UTR + CP partial	17.6x	180	2050
RNA2 MP + CP	9.7x	156	3344

(b) ONV1 sequences generated by concatenating and then assembling the file containing reads from the a pooled sample (PL-BLK) with the higher depth, single-sample dataset reads.

Segment	Depth (average)	Reads used	Length
RNA1	21.1x	831	7663
RNA2	15.9x	460	5902

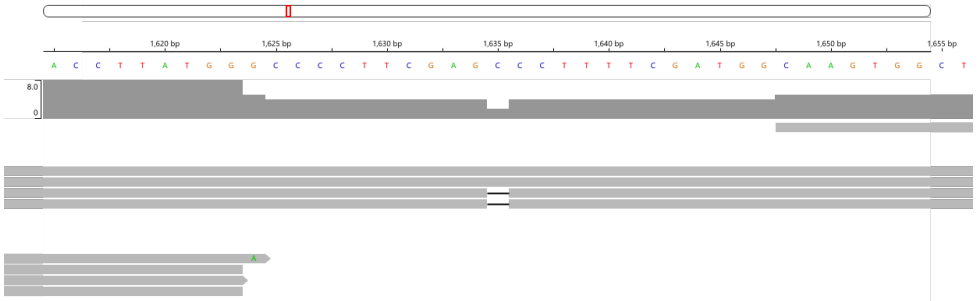


Figure 3.24 A view of the reads from the concatenated PL-BLK/PL24 datasets mapped to the region of ONV1 RNA2 affected by the indel, generated by IGV Web.

The ArMV detected infecting the oca plants obtained from Europe was not detected by the ELISA using the ArMV antisera kept at Fera or by the qRT-PCR using unpublished ArMV primers from Fera Science Ltd.. The latter case may be due to difficulties with the buffer, which is proprietary and no longer being manufactured in any case. The failure to detect the oca isolate of ArMV using commercial antisera might suggest that it is a biologically distinct isolate with different serological properties from the type strain.

Sequence likely representing a divergent isolate of Arabis mosaic virus large satellite RNA was detected in the FR-BLK dataset. Eight nepoviruses have large satellite nucleic acids associated with them, two each from subgroups A and B, and four from subgroup C. The genus *Nepovirus* also contains three species which can be associated with small satellite RNAs, two from subgroup A and one from subgroup C (King et al., 2011). In the case of Arabis mosaic virus, the effects of co-infection with Arabis mosaic virus large satellite RNA depend in the family of the host; in some cases symptoms are worsened, in others ameliorated (LIU et al., 1991). Without a comparison in oca or the wider *Oxaliaceae* family, it remains to be shown whether the satellite in these plants is having any effect on the symptoms.

Oca nepovirus 1 From the BLASTN and BLASTX analyses, three sequences in the PL-BLK dataset were identified that had distant (55% amino acid identity) homology to cherry leaf roll

virus. These sequences varied in length from 2050 to 5000+ base pairs. Two were identified by Trinity as isoforms of the same sequence, which indicates that parts of the sequences were considered by the software to be similar enough in part of the assembled sequence to belong to the same gene family. This suggested that the similar sequence could represent one larger scaffold if aligned. An alignment of these two isoforms in MEGA X created an overlap that resulted in a longer scaffolded sequence representing the majority of ORF2 and part of the 3' UTR of RNA2. Based on the distant homology to cherry leaf roll virus, this putative novel virus was named oca nepovirus 1 (ONV1).

In order to obtain more of the genome of ONV1, total RNA was extracted from tissue of one plant that tested positive via PCR, PL24. This RNA was then sequenced by HTS at a higher depth. The dataset generated contained one 692bp sequence with distant homology to cherry leaf roll virus, matching only a small portion of the 3' UTR (27% query cover, accession KF779205.1). This was likely to be sequence of ONV1. The recovered sequence was only one nucleotide different from the ONV1 3' UTR of the RNA2 sequence derived from from the PL-BLK dataset. To determine if more sequence matching to ONV1 had been sequenced but not assembled, the reads were mapped with BBMap to the ONV1 sequence obtained from the PL-BLK HTS dataset. The percent coverage was 50.94%, suggesting that further ONV1 sequence had been detected by the machine, but this was fragmented and thus could not be assembled by Trinity. Based on this finding the reads from the PL-BLK and PL24 datasets were concatenated into one file and processed together using Angua. This allowed considerably more RNA1 to be recovered, and a further portion of the RNA2, compared with the original pooled dataset. In addition, the RNA2 sequence was assembled as one sequence rather than two. The average coverage for RNA1 was slightly higher (21.1x instead of 18.9x), but for RNA2 reflected a coverage in between the coverage of the initial two RNA2 sequences from the PL-BLK dataset (15.9x instead of 17.6x and 9.7x respectively).

The RNA2 sequence recovered from HTS suggested a stop codon at nt 1120 of the original sequence (corresponding to nt 1635 of the later, more complete sequence), which is premature for nepoviruses (Fuchs et al. 2022). To resolve whether this stop codon was an assembly error, primers were designed to the region between nt 589 and nt 1825 on the original PL-BLK dataset sequence (corresponding to nt 1106 to nt 2340 of the more complete sequence). These primers were used to amplify extracts from plants which had tested positive for ONV1 using qRT-PCR as well as a pooled extract containing all of the plants that were extracted and sequenced as the PL-BLK dataset. These were samples PL5, PL24, PL25, PL27, and the pooled extract. The extracts from individual plants were subjected to PCR as a 1 in 10 dilution and as neat extracts. Sanger sequencing of the amplified sequence extracted from the visualisation gel provided a sequence containing an additional cysteine residue, restoring the polyprotein ORF and bringing the known ONV1 sequence to 5903 base pairs.

By mapping the reads from the concatenated file to the assembled sequences, the reads used to assemble the sequence could be analysed. The region which, in the sequence, contained the stop region had only four reads mapped to that site. Two of those four reads lacked the cysteine

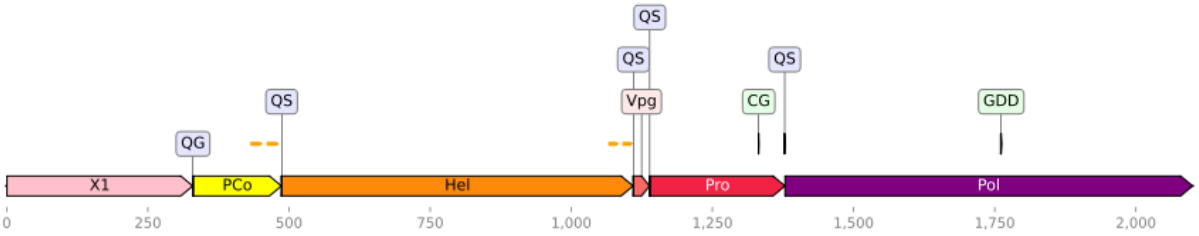


Figure 3.25 Putative cleavage sites for ONV1's RNA1 polyprotein. Note that X1 is likely to continue in the N terminal as the sequence is not coding complete. The small orange arrows above the main domains indicate predicated transmembrane helices. The conserved *Secoviridae* CG and GDD ProPol motifs are also highlighted.

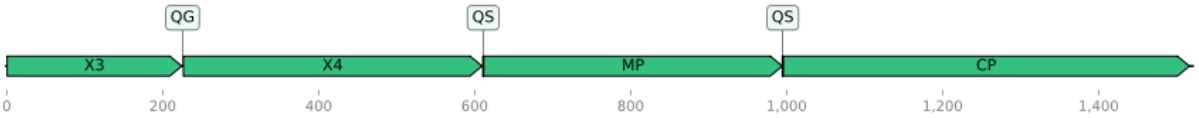


Figure 3.26 Putative cleavage sites for ONV1's RNA2. Note that X3 is likely to continue as this segment is unlikely to be coding complete.

residue, while two had the correct residue (Figure 3.24). In assemblies from both the PL-BLK and PL24 HTS reads, the Trinity assembler produced a sequence with the cysteine absent. This error was only resolved by Sanger sequencing.

Nepoviruses express their genes as two polyproteins which are then cleaved by a 3C-like cysteine protease to the mature products (Rott et al., 1995). By comparing the putative novel virus with cherry leaf roll virus accessions and other sequence belonging to members of the genus *Nepovirus*, it is possible to tentatively place the position of the cleavage sites. As seen in the phylogeny (paragraph 3.2.11), the ONV1 sequence clusters with CLRV, tomato ringspot virus, and other members of the C-1 clade (Sanfaçon, 2022). Subgroup C nepoviruses are notable for the size and arrangement of their RNA1, which is generally longer than the other two subgroups and has X3 and X4 genes before the MP and CP genes, leading to an extra cleavage site (Fuchs et al., 2017).

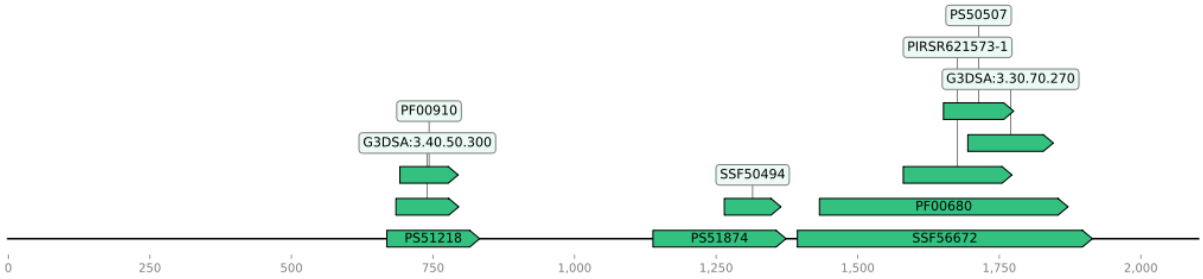


Figure 3.27 A summary of the InterProScan domains for the ONV1 RNA1. Broadly the three groups are: helicase superfamily 3 hits, Picornavirales 3C-like protease hits, and ssRNA viral RdRp hits.

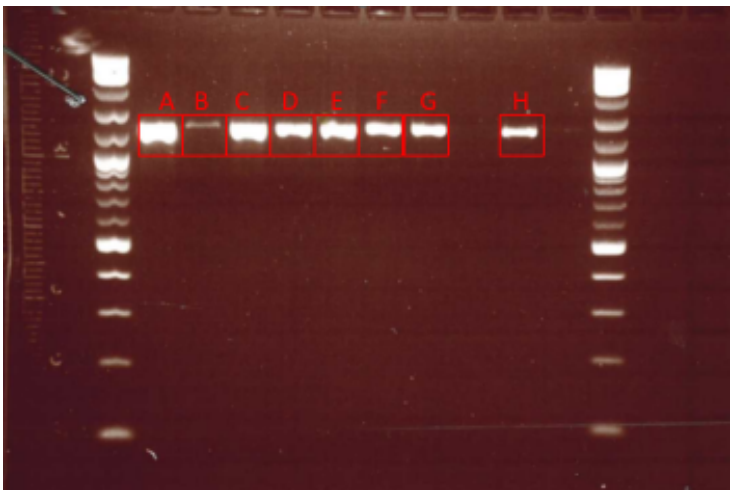


Figure 3.28 PCR products used for Sanger sequencing of oca nepovirus 1. A = PL05, B= Pl24, C=PL25, D=PL27, E=Pooled extracts, F=PL05 diluted, G=PL24 diluted, G=PL27 diluted. The dilution of PL25 did not produce a band when visualised.

In the case of CLRV, Von Bargen et al. in 2012 found that only the CP domain of RNA1 matched any accessions on the CDD-database (Von Bargen et al., 2012). This is the case for ONV1 as well using the InterProScan database. A scan using InterProScan (Jones et al., 2014) found matches to the N terminal, 'coat protein' and C terminal of nepoviruses, but no other viral domains were found for this segment. On RNA1, other domains were found, which helped inform the tentative placements of the cleavage sites. A summary of domains found by InterPro and their placement is shown in Figure 3.27 on Page 82.

Based on the presence of a transmembrane helix and highly hydrophobic portion of the gene in the putative PCo region of CLRV (Von Bargen et al., 2012), the X1/PCo cleavage site has been tentatively placed at 330aa, at a Q/G site. This preserves a transmembrane helix at the C terminal and creates a highly hydrophobic protein of predicted 16.69 kDa. Q/G then Q/S for the first two proteins was reported for another C1 nepovirus, babaco nepovirus 1 (Sanfaçon, 2022).

The presence of the novel nepovirus in individual plants was determined by amplifying extracts with qRT-PCR primers designed from the sequences. In total 4 out of 27 plants that formed the PL-BLK pool tested positive for ONV1.

During the specificity test, none of the extracts used as positive controls tested positive using the primers designed to ONV1, including the CLRV control from DSMZ. Testing of further plants purchased from the internet gave negative results when tested using these primers (see subsection 3.2.15).

The cnPCR and qRT-PCR primers were designed to RNA2 and RNA1 respectively. Both sets of primers gave positives from the same four plants as well as the pooled extract. This further suggests that the two segments found were indeed from the same virus.

A cutting taken from an infected plant from samples forming the PL-BLK pooled sample dataset tested positive qRT-PCR for ONV1 with assayed with qRT-PCR, confirming a systemic infection.

Extracts from ONV1 positive plants did not cross-react with CLRV antisera (DSMZ, e and ch strains).
When amplified using generic nepovirus subgroup C primers presented by Digiaro et al. in 2007 (Digiaro et al., 2007), all CLRV nucleotide extracts (DSMZ, e and ch) tested negative with all master mixes and cycling conditions attempted. Based on this, ONV1 was not tested with these primers to determine if it would be amplified as a subgroup C nepovirus.

Oca nepovirus 2 Sequences (13 in total) were recovered from the 2019 dataset which matched distantly to blueberry leaf mottle virus (BLMoV), with 59.4-71% amino acid identity. These sequences were previously annotated as BLMoV in the data provided by Ian Adams (Fera Science Ltd.). Based on the homology to nepoviruses, this virus was named as oca nepovirus 2 (ONV2). Analysis of the sequences from the datasets in which ONV2 (or BLMoV) was detected revealed that both RNA1 and RNA2 of ONV2 were sequenced from extracts of several samples purchased and investigated for viral infection in 2019. One small sequence was recovered from sequencing of newly extracted RNA of plants which were retained for several years after the initial study in 2019. The 5' UTR of the sequences deemed to represent RNA1 and RNA2 is nearly identical, suggesting that these segments belong to the same species.
The proteinase cofactor domain, usually split into X1-X2(PCo), lacked the N/G cleavage site proposed as shared between all clade C-3 nepoviruses (Sanfaçon, 2022). The X2 protein is a transmembrane protein (Zhang and Sanfaçon, 2006). Aligning all three C-3 clade members as reported by this paper, the N/G site close to a predicted TM Helix was conserved between all three other members. This site was D/G in the novel nepovirus. Since D/G is a reported (tentative) cleavage site for the related grapevine Bulgarian latent virus, it seems plausible that this is the cleavage site for this isolate, especially with the alanine in the P3 position (Sanfaçon, 2022). Thus the D/G cleavage site is tentatively proposed here. Accession AF368272 of BRV is annotated with a mature chain ending at this N/G site before beginning another directly after as well. ONV2 is likely to have the same cleavage sites as BRV in the other RNA1 sites: NA for X2/NTB (or PCo/Hel) and NTB/Vpg (GALSAN/A, though this interrupts a predicted transmembrane region), and NS for Pro/Pol (Sanfaçon, 2022), with the exception of NA instead of NS between VpG and Pro, which isn't reported at this site in the other clade C-3 nepoviruses. However, the presence of the semi-conserved VpG motif in the putative VpG product suggests that the cleavage site is here (Mayoa and Fritsch, 1994), and there is no corresponding N/S cleavage site within the correct region.
This nepovirus contains the conserved protease cofactor (F-x27-W-x11-L-x21-L-x1-E) motif experimentally confirmed in members of some *Nepovirus* species (Pacot-Hiriart et al., 2001), with the exception of a leucine at position 512, though accessions YP_004429248 (grapevine Bulgarian latent virus) and YP_010840563 (blueberry leaf mottle virus) also didnot have a leucine at this position of the conserved motif when aligned in Geneious Prime (Figure 3.29). The conserved protease cofactor motif sits within the predicted PCo region. The putative organisation of the RNA1 segment is shown in Figure 3.30.

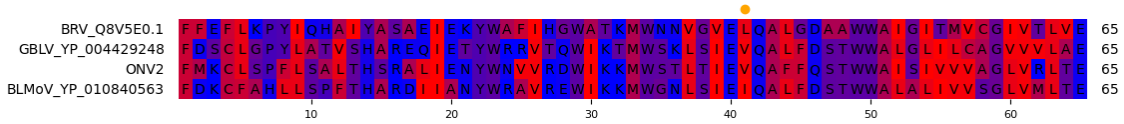


Figure 3.29 Alignment of the protease cofactor region showing the leucine or other residue in each member of clade C-3.

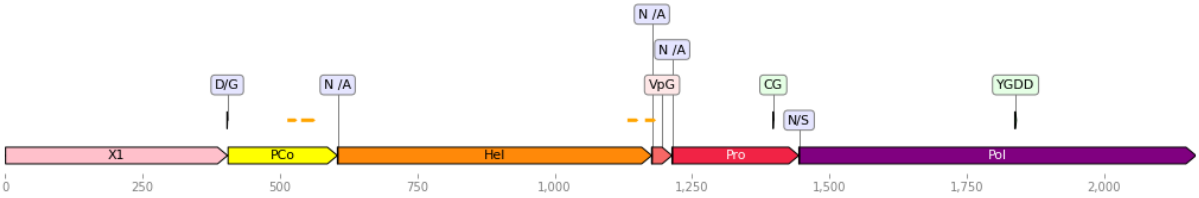


Figure 3.30 Putative cleavage sites for the RNA1 polyprotein (P1) of oca nepovirus 2.

The cleavage sites of RNA2 are the same as those predicted for BlMoV; based on alignment, all of the mature products cleavage at an N/S position (Sanfaçon, 2022). Based on the findings of Sanfaçon, the consensus sequence for BlMoV/GBLV/ONV2 for the X3/X4 cleavage site is GDXTCN/S, for all C-3 nepoviruses is GDLTCN/S at the X4/MP site, and for all C-3 nepoviruses is RXXTCX/S at the MP/CP site. ONV2 differs from the other viruses in this subgroup for not having an phenylalanine at the -5 position on this last cleavage site.
The ProPol is 80.7% similar in amino acid sequence to blueberry leaf mottle virus. The coat protein is 41.4% identical in amino acid sequence to blackcurrant reversion virus.
A search on Serratus found no same-species hits, nor any SRA entries containing reads with identity greater than 90% to that of ONV2.

Phylogeny of Nepoviruses To place them within the genus, both novel nepoviruses discovered in oca were placed on a phylogenetic tree together. The ProPol region is highly conserved in the *Secoviridae* family and is useful for inferring phylogeny (Sanfaçon et al., 2009), as is the sequence representing the coat protein in nepoviruses, despite only representing a small piece of the genome (Steinkellner et al., 1992).
Phylogenetic analysis of the ProPol region of exemplar isolates confirmed the close relationship of ONV1 to cherry leaf roll virus (CLRV) and Stenotaphrum nepovirus, and ONV2 to grapevine Bulgarian latent virus (GBLV) and BLMoV (Figure 3.32). This relationship is supported by the

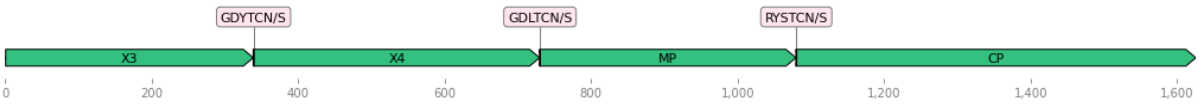


Figure 3.31 Putative cleavage sites for the RNA2 polyprotein (P2) of oca nepovirus 2.

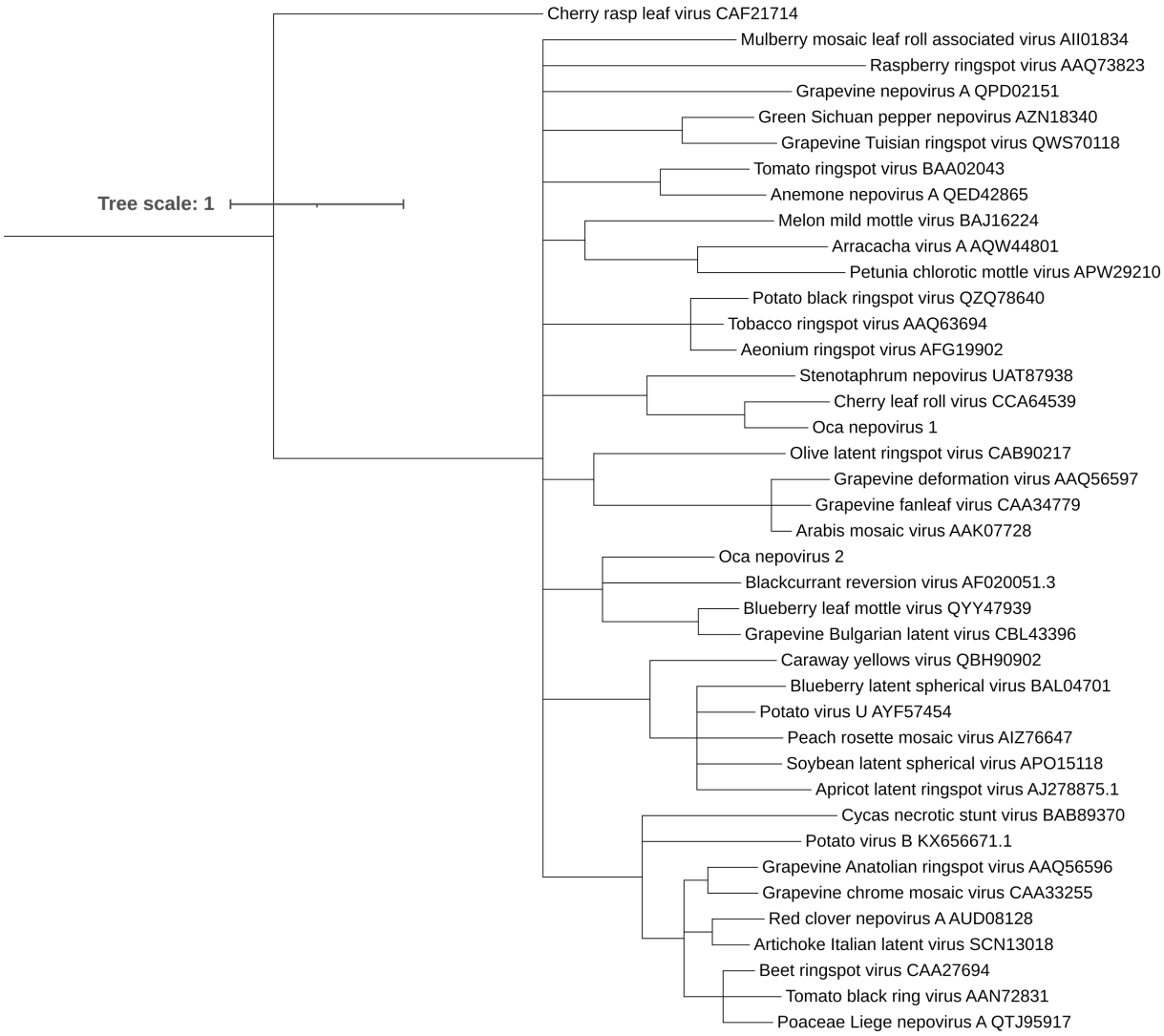


Figure 3.32 Phylogenetic tree of both novel nepoviruses based on the amino acid sequence of the putative Pro-Pol conserved sequence.

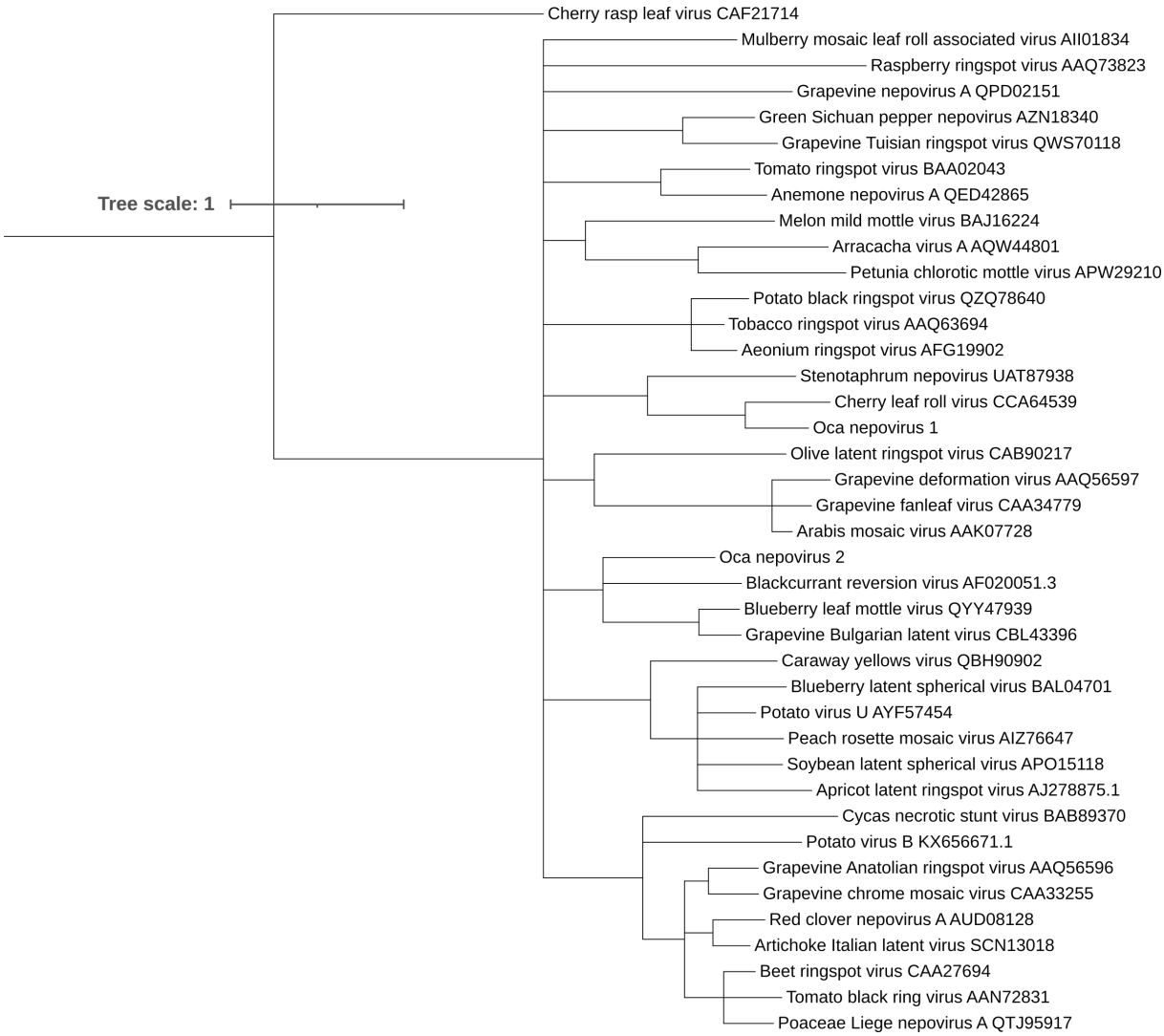


Figure 3.33 Phylogenetic tree of both novel nepoviruses based on the amino acid sequence of the putative coat protein.

assignment of the ONV2 sequences to either GBLV or BLMoV by the MEGAN software based on homology.
The ICTV species demarcation threshold for nepoviruses is less than 75% amino acid identity in the coat protein or greater than 80% amino acid in the Pro-Pol region (Fuchs et al., 2022). Though ONV2 is on the edge of this threshold in the ProPol region at 80.7% identity, it is well below this threshold in the coat protein. Recombination analysis would be required to determine if the ONV2 sequence here is a recombinant between a highly divergent BLMoV isolate and ONV2 or if this sequence is representative of ONV2 as a whole. The sequence of ONV1 is clearly below the demarcation threshold in both genes. The nepoviruses presented here are suggested as members of two new species, *Nepovirus primoxalidis* (oca nepovirus 1, ONV1) and *Nepovirus secundoxalidis* (oca nepovirus 2, ONV2). Based on genome organisation (X3-X4 domains in the RNA2 (Fuchs et al., 2017)) and phylogeny, both viruses are likely to be members of subgroup C.

3.2.12. Oca alphacytorhabdovirus 1

A number of sequences from the HTS datasets showed distant homology (46-74% amino acid identity when translated with BLASTX) to accessions of the cytorhabdoviruses Medicago alphacytorhabdovirus 1, Artemesia alphacytorhabvirus 3, alfalfa dwarf virus and Hyptis latent virus. This possible novel virus is referred to as oca alphacytorhabvirus 1 (OACV1). One sequence, from the L20S11 HTS dataset, contained a predicted open reading frame that translated a 2095 amino acid product matching the replicase (L gene) ORF. This suggests that the entire replicase of this novel virus was recovered, allowing a phylogeny even without a complete genomic sequence. Additionally, a search of the predicted amino acid product on Serratus found genus-level (70-90%) homology to reads originating from transcriptomes of *Medicago sativa* and to NCBI isolates of alfalfa dwarf virus from *M. sativa* in Argentina.
Other sequences when searched via BLASTX showed homology to the nucleocapsid (N) protein, glycoprotein (G), movement (M), phosphoprotein, P3 and P6 proteins. These portions shared similar relationships to other cytorhabdoviruses as the replicase sequence did. The longest single sequence was 8336 base pairs long, a sequence containing the L/replicase ORF. The replicase is 58.8% similar in amino acid sequence to that of strawberry crinkle virus. The longest recovered sequence had 54.9% amino acid identity to strawberry crinkle virus across the represented region.
The phylogeny was constructed using ratified members of the current genus *Cytorhabovirus*, and a larger one with these viruses and also proposed members of a new genus, *Alphacytorhabdovirus* (Bejerman et al. 2023). The replicase sequence of OACV1 clusters (and is thus likely to share its most recent common ancestor) with strawberry crinkle virus, alfalfa mosaic virus, Hyptis latent virus and raspberry vein chlorosis virus. Strawberry crinkle virus, responsible for severe symptoms in strawberry (*Fragaria* × *ananassa*, family *Rosaceae*), and geographic origin does not seem to reflect in the phylogeny of isolates (Klerks et al., 2004). It is widespread, found all over the world, including rare detections in Argentina, Chile and Brazil

(CABI, 2024). Alfalfa dwarf virus, which also causes economically relevant symptoms, is endemic to Argentina, showing very little variation in its N gene between geographic isolates, and its primary host is alfalfa (*Medicago sativa*, *Fabaceae*) (Samarfard et al., 2018). Raspberry vein chlorosis virus has at least two very genetically distinct isolates (only 68% nucleotide similarity to one another), primarily infects raspberry (*Rubus* subgen. Idaeobatus, *Rosaceae* (Graham et al., 2007)) symptomatically and has a wide distribution (Jones et al., 2019b). Hyptis latent virus is a recently discovered virus only detected at present in Ecuador, found on wild *Hyptis pectinata* (*Lamiaceae*) plants and found to be vectored by the pea aphid (*Acyrthosiphon pisum)*, with no symptoms demonstrated (Reyes-Proaño et al., 2022).
Based on the disparity of the sequences (representing non-overlapping regions of the cytorhabdovirus genome) and the fact that the sequences originate from different samples in most cases, it is not possible to say for certain that the sequences belong to only one species. The sequence presented here as partial sequence of a novel cytorhabdovirus is the replicase only, with the most parsimonious interpretation being that the other cytorhabdovirus sequences from different genes that were detected by sequencing RNA extracted from the oca samples all belong to the same species. Genomic regions represented in multiple sequences were within the species demarcation threshold when aligned with one another.
The ICTV species demarcation critera for cytorhabdoviruses require the full genome in theory, but suggest 80% similarity or less in all open reading frames, or complete genomes with 75% similarity or less (Walker et al., 2022). Based on these findings, this replicase gene is likely to belong to a novel species of cytorhabdovirus, specifically the proposed genus *Alphacytorhabdovirus* (Bejerman et al., 2023). This partial sequence is suggested under the name *Alphacytorhabdovirus oxalidis* (oca alphacytorhabdovirus 1, OACRV1).

3.2.13. Novel Virus Fragments

The L21S5 HTS dataset also contained sequence which might correspond to RNA1 of a novel emaravirus distantly homologous to the coat protein of pigeonpea sterility mosaic emaravirus 2 (76.6%). Only the middle portion of the sequence matched any accessions, making for 49.38% coverage.
The sequence data (L22S3) also contained a genome of a possible novel deltaflexivirus, whose main ORF (replicase) was 85% identical in nucleotide sequence to Xinjiang deltaflexi-like virus. Deltaflexiviruses are usually mycoviruses (Bejerman and Debat, 2022).

3.2.14. qRT-PCR of Individual Plants

Using primers designed to the virus sequences detected in the pooled HTS datasets (except UPolV1, which used primers from the paper first reporting this virus (Fox et al., 2019), the individual plants whose extracts were pooled to make each HTS dataset were re-extracted and tested for presence of the viruses using qRT-PCR. In total, 4/9 of the plants whose extracts were pooled to form the FR-BLK dataset were infected with the viruses tested for. Of these four plants which tested positive for at least one virus, all were infected with more than one virus

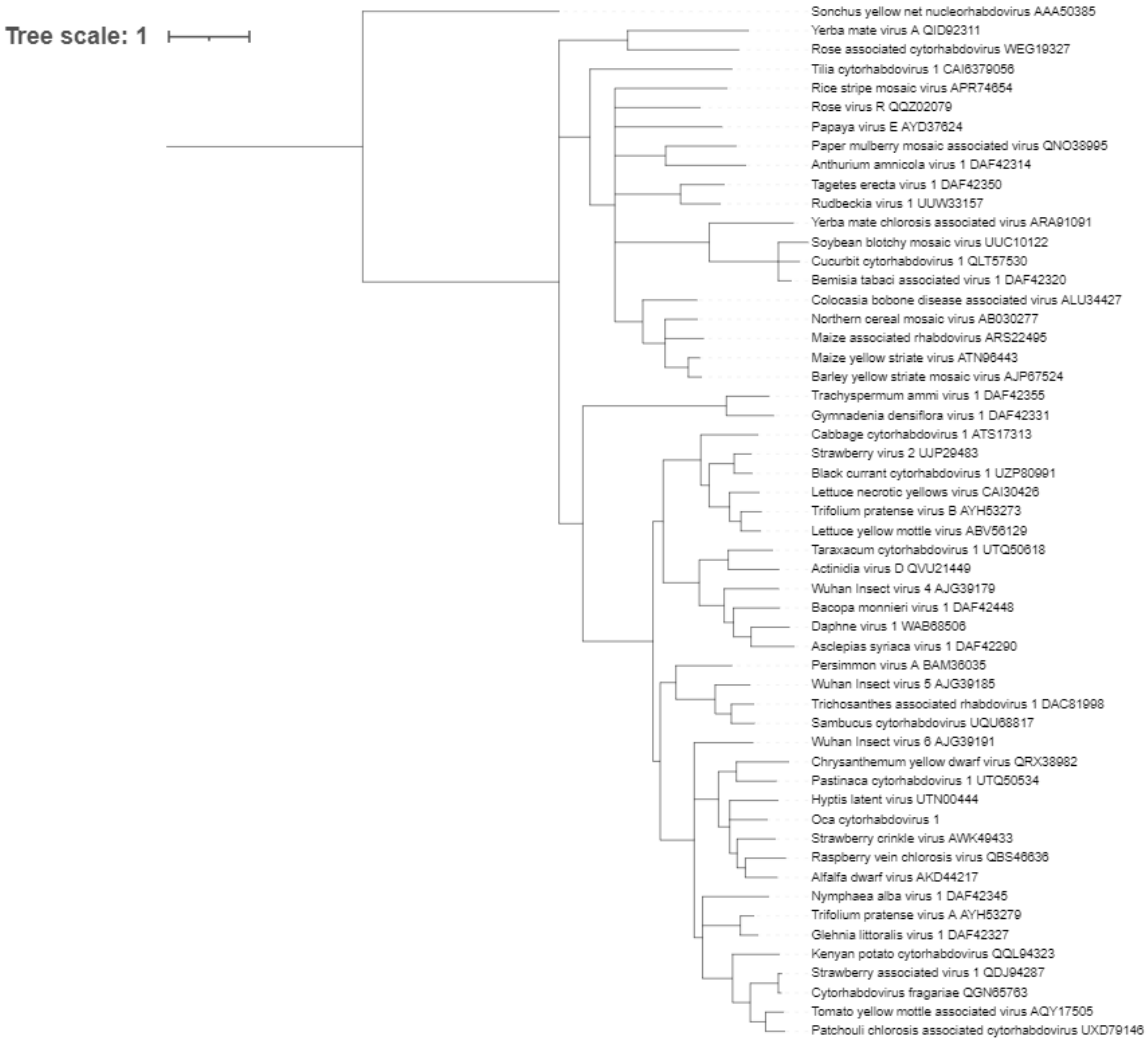


Figure 3.34 Phylogenetic tree of all ratified members of the current genus *Cytorhabdovirus.*

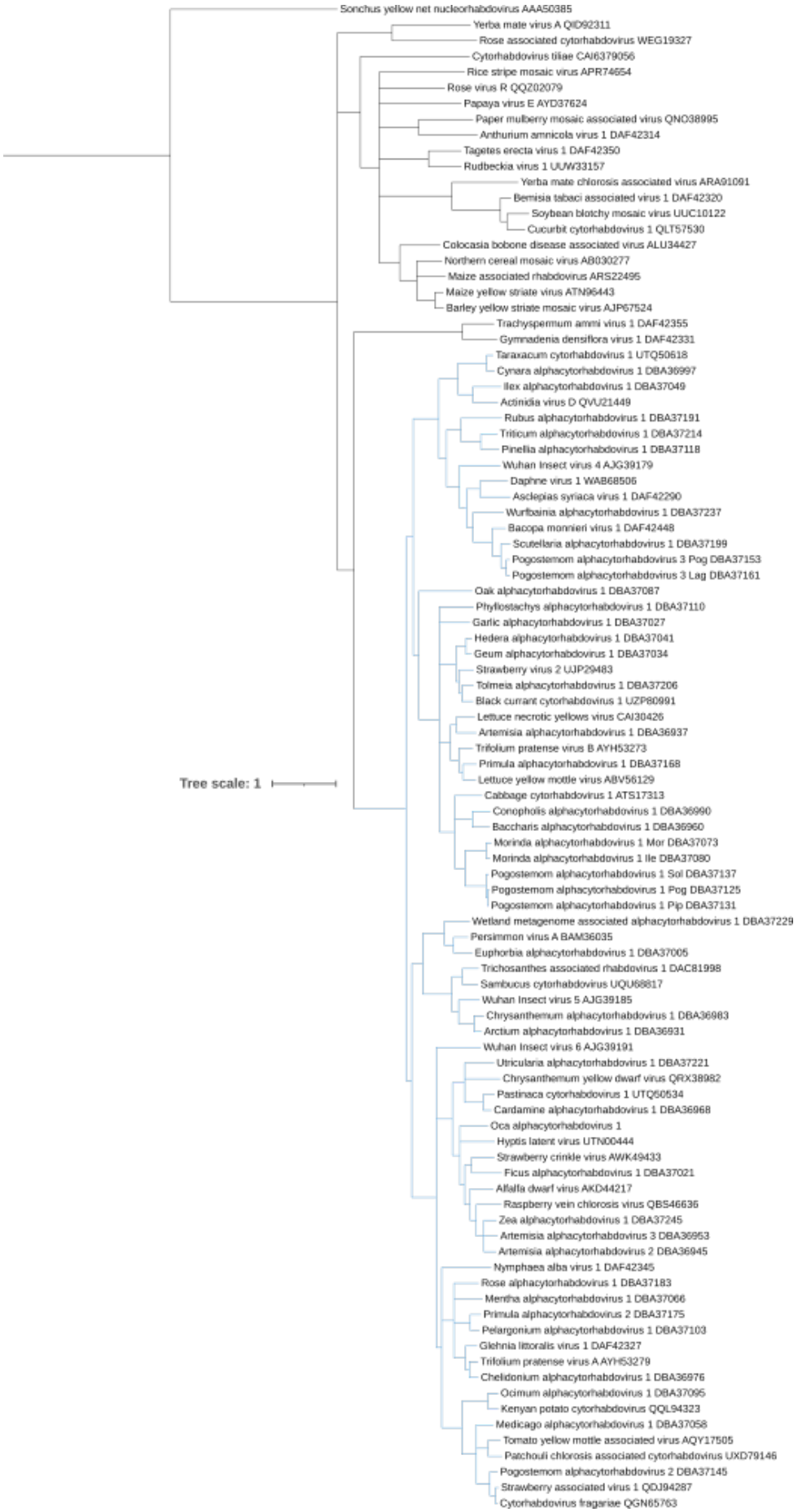


Figure 3.35 Phylogenetic tree of all ratified members of *Cytorhabdovirus* plus all proposed members of *Alphacytorhabdovirus* (marked in blue).

Table 3.4 A table of the novel viruses detected by qRT-PCR within the PL-BLK and FR-BLK datasets and reads generated from sequencing of RNA extracted from individual samples, including tentative detections of at least one novel ophiovirus using the generic primers. OcaV1 was not tested for using qRT-PCR.

Sample	In Pool	ONV1	ArMV	OAV1	OVX	UlPolV1	Oph.	OCV1
PL01	Poland			Y		✓	✓	
PL02	Poland							
PL03	Poland							
PL04	Poland					✓		
PL05	Poland	✓						
PL06	Poland			✓		✓		
PL07	Poland		✓					
PL08	Poland					✓		
PL09	Poland							
PL10	Poland					✓		
PL11	Poland							
PL12	Poland							
PL13	Poland			✓		✓		
PL14	Poland		✓					
PL15	Poland					✓		
PL16	Poland							
PL17	Poland							
PL18	Poland							
PL19	Poland							
PL20	Poland						✓	
PL21	Poland					✓		
PL22	Poland							
PL23	Poland							
PL24	Poland	✓	✓					
PL25	Poland	✓						
PL26	Poland							
PL27	Poland	✓						
LY	France							
OB	France		✓					✓
DK2	France		✓		✓			
DK1	France		✓		✓			
B1	Poland		✓	✓	✓			✓
Amar	France							
An	France							
OAEC	France							
WP	France							

tested for. One plant, B1, was determined to be infected with four separate viruses. 15/27 plants extracted and pooled together as PL-BLK tested positive for infection with at least one of the viruses tested for. Five of the plants were found to be infected with more than one of the viruses tested for. That equates to one third of the infections being mixed.
The primers designed to the ophiovirus sequence obtained from the FR-BLK dataset, produced inconsistent results, with late Ct values or negatives occuring as frequently as strong positives when the same extract was tested multiple times. This may represent a need for validation of primers, or merely very low titre or inconsistent distribution within the host cells. The Hou et al. paper in 2020 which first pointed out the dearth of biological information accompanying HTS virus reports doesn't specify how many of the primers reported in the papers that they investigated were validated (Hou et al., 2020). Generally, it is quite easy to design primers using dedicated software, but ensuring that the primers have correct melting temperatures to ensure adequate specificity and sensitivity is time-consuming (Korbie and Mattick, 2008).
This meant that the viruses OVX, OAV1 and OCV1 were only present in mixed infections in the plants sampled. Those likely to be singly infected were infected with either ONV1 (3), ArMV (2), and UlPolV1 (4). In total, 17/37 of the original plants had a negative result for all viruses tested for, with a caveat: the ophiovirus primers were not deemed reliable. This likely reflects the low recovery of the genome from the pooled sample dataset PL-BLK, which provided only one small sequence to which the primers were designed.
One plant (PL24) was sampled and the total RNA extract subjected to further HTS. This sequencing suggested that the plant was virus free, but very low read numbers cannot preclude the possibility of a low titre virus being missed. These low read numbers could, however, indicate successful host RNA subtraction.
In addition to the qRT-PCR tests, OVX was detected in 3 plants by amplifying using primers designed to detect all members of the genus *Potexvirus* (Van Der Vlugt and Berendsen, 2002). The detection of OVX with potexvirus-specific primers gives further evidence that this virus is a likely new member of the genus *Potexvirus*.
In terms of ONV1, amplification using conventional RT-PCR primers designed to the RNA2 as well as the qRT-PCR primers designed to the RNA1 both provided positive results on extracts from the same samples. This amplification of ONV1 sequence in extracts from the same plants by different primers in different formats helps lower the chance of the reported genome being chimeric in origin, i.e. associating an RNA1 and an RNA2 sequence from two different nepoviruses to one another erroneously. ONV1 appears to have a low incidence. Of 75 plants tested from several sources, 4 tested positive for ONV1. All of the plants which tested positive for ONV1 which were plants whose RNA was extracted and pooled to form the same HTS dataset in which ONV1 was first detected (PL-BLK). This virus was not detected by a qRT-PCR assay of any additional samples, even other plants purchased from other sellers in Poland. This may suggest that ONV1 is not widespread even in the country where it was first detected, but low sample sizes make this impossible to say for certain. It could be that ONV1 as a population is South American in origin and that the ONV1-infected plants purchased from Poland descend

from plants that were infected with ONV1 in South America. ONV1 could also be an indigenous European virus which is not widespread. Finally, it could be widespread in Europe, but oca could be a host which ONV1 does not readily infect.
The HTS and qRT-PCR survey data suggest that OAV1 in particular is very widespread in oca plants in Europe. Samples from several countries were infected. Whether this reflects a biosecurity concern, i.e. a pathway which is disseminating an introduced virus widely, or a detection of a common, native European allexivirus is again uncertain. Ongoing baseline studies of managed or unmanaged European flora could help to determine if this virus is emerging onto introduced oca from European plants if it is detected widely in European flora (e.g. (Maclot et al., 2020)). OCV1 is likely already widespread in the UK given that it was detected in weeds that were not adjacent to any allotments.

3.2.15. Survey of Further Internet Purchases

Further samples were purchased from the internet in order to survey for four of the novel viruses detected in the original HTS datasets (PL-BLK and FR-BLK).
In total, 39 additional plants were subjected to RNA extraction and qRT-PCR. These additional samples originated from the UK, Hungary, Ireland, the Netherlands and Poland, making for 75 plants in total which were extracted and tested with qRT-PCR. This survey focused on four viruses described in this thesis, detected in the FR-BLK and PL-BLK datasets: OVX, ONV1, OCV1 and OAV1.
66.6% (two thirds) of the oca plants were determined to be infected with at least one of the four viruses tested for. Generally, mixed infections between these novel viruses were common; 17.9% of plant samples tested positive for mixed infections, increasing to 28% of plants infected with any virus being infected with a mixed infection. The true number is likely higher when accounting for Arabis mosaic virus and UlPolV1 infections, which were common in the HTS datasets but were not tested for in the subsequent purchases.
Slightly over half of the oca plants (53.8%) tested were found to be infected with OAV1. In two purchases of three or more plants, every plant tested positive for OAV1. Importantly, OAV1 was detected in leaf tissue from the registered variety 'Giggles'. ONV1 infections, in contrast, were detected rarely in the 27 plants whose extracts were pooled together as PL-BLK. ONV1 was not detected in any plants from later purchases, even from Poland, nor in plants extracted as part of the 2019 HTS datasets which originated from Poland. OVX was also detected at a relatively low incidence, detected in only 5% of extracts, two of which came from plants from the same purchase. OCV1 was detected in 23% of samples tested, including 3/4 from a purchase from one commercial supplier and every extract of 'rainbow mix' oca plants from the seller in Ireland. No plants from a purchase from a garden centre or other commercial supplier was determined to be virus free.
Only two purchases consisted only of plants deemed negative for all of the viruses tested for, that being the 'oca de perou' purchase from the UK and the 'White' purchase from Poland.

Table 3.5 A table of findings of the four novel viruses tested for in the survey samples.

Sample	Variety	Origin	ONV1	OAV1	OCV1	OVX	Tissue
WW1	White Wonder	Hungary		✓			Leaf
WW2	White Wonder	Hungary					Leaf
WW3	White Wonder	Hungary					Leaf
Rf1	“Red”	France		✓			Tuber
Rf2	“Red”	France		✓			Tuber
Rf3	“Red”	France		✓			Tuber
Rf4	“Red”	France		✓			Tuber
Rf5	“Red”	France		✓			Tuber
MB	Mashua blanca	Ireland					Leaf
RM1	“Rainbow mix”	Ireland			✓		Leaf
RM2	“Rainbow mix”	Ireland			✓		Leaf
RM3	“Rainbow mix”	Ireland		✓	✓		Leaf
RM4	“Rainbow mix”	Ireland		✓	✓	✓	Leaf
RM5	“Rainbow mix”	Ireland		✓	✓	✓	Leaf
W1	“White”	Poland					Leaf
W2	“White”	Poland					Leaf
Ruk1	“Red”	UK		✓			Tuber
Ruk2	“Red”	UK		✓			Tuber
Ruk3	“Red”	UK		✓	✓		Tuber
Ruk4	“Red”	UK		✓	✓		Tuber
Gig	Giggles	UK		✓			Plant
Stem	Unknown	UK					Stem
Rcom1	“Red”, commercial	UK		✓	✓		Leaf
Rcom2	“Red”, commercial	UK			✓		Leaf
Ycom1	“Yellow”, commercial	UK			✓		Leaf
Ycom2	“Yellow”, commercial	UK		✓			Leaf
Rcos	“Red”, ’cosmetics’	UK		✓			Tuber
NZ	New Zealand	UK		✓			Tuber
OdP1	Oca de Perou	UK					Tuber
OdP2	Oca de Perou	UK					Tuber
OdP3	Oca de Perou	UK					Tuber
OdP4	Oca de Perou	UK					Tuber
OdP5	Oca de Perou	UK					Tuber
OdP2	Oca de Perou	UK					Leaf
OdP3	Oca de Perou	UK					Leaf
OdP5	Oca de Perou	UK					Leaf
C1	“Cream”	UK		✓			Tuber
C2	“Cream”	UK		✓			Tuber
C3	“Cream”	UK		✓			Tuber
C4	“Cream”	UK					Tuber
C5	“Cream”	UK					Tuber
C6	“Cream”	UK					Tuber
C7	“Cream”	UK					Tuber

None of the samples of RHS accessions tested positive for any of the novel virus candidates tested, including OCV1 which was detected in a small number (one fifth) of UK *Oxalis* extracts (1/3 wild and 1/7 cultivated, 2/10 total).

3.3. Discussion

For this chapter the virome of oca available for purchase in the United Kingdom was investigated. As well as two known viruses (ArMV and UlPolV1), a number of novel viruses in different families were detected through HTS. Most of the sequences found representing putative novel viruses are complete or near-complete genomes. In addition, four of these viruses were tested for in a further set of internet-traded samples using qRT-PCR.
The plants investigated here represent the types of material commonly found for sale on sites such as eBay through informal sellers, but also some plants available from larger, commercial nurseries via internet trade.
The CMV sequence detected here was considered to be contamination, as CMV was not detected by qRT-PCR or ELISA in any of the individual plants sampled. Given that the UDI does not entirely eliminate cross-talk, 'index hopping' is also possible where this sequence may have been mis-assigned by the base call software (Macconaill et al., 2018). Similarly, the L20S11 dataset contained sequences matching to African accessions of maize yellow dwarf virus. An investigation of datasets from other samples sequenced on the run found this virus in datasets originating from sequencing of RNA extracted from maize, and the sequences likely originate from contamination due to these infected maize plants (data not shown).

3.4. Conclusion

The viruses detected in these oca samples belong to completely different viral taxa with different lifestyles and niches. Ophioviruses are negative-sense single-stranded RNA viruses, that are filamentous and multipartite (Vaira and Milne, 2008). Potexviruses and allexiviruses are both from the family *Alphaflexiviridae*, flexuous filamentous positive-sense single-stranded RNA viruses (Kreuze et al., 2020b). Poleroviruses form isometric particles, containing one positive-sense single-stranded genomic segment and (Mayo and Ziegler-Graff, 1996). Nepoviruses are single-stranded, positive-sense RNA viruses with two segments encapsidated separately in icosahedral, sphere-like particles (Hull, 2001).
Included in this virome are plausibly two DNA viruses as well: a capulavirus, which has an unsegmented circular ssDNA genome (Varsani et al., 2017), and a caulimovirus, dsDNA viruses with isometric particles (Hull, 2001). Should either of these be endogenous, it would indicate an infection in the recent or distant past.
Other members of the genera to which the viruses presented here likely belong are thought to contain members vectored by pollen, seed, nematodes and mites (*Nepovirus* (Thresh, 1964, Isogai et al., 2020, Lister and Murant, 1967)) by fungi (*Ophiovirus*, (Viara and Hammond, 2020)) by aphids, leafhoppers and other insects (Ryckebusch et al., 2022, Teycheney et al.,

2020) (Several), and by contact (*Potexvirus* (Verchot, 2022)). In other words, there is an astonishing variety of viral lifestyles represented by the viruses detected in these oca plants. Perhaps most interesting is the abundance of viruses which do not necessarily rely on aphid transmission. Though the vectors of these specific viruses are unknown and indeed the vectors of *all* non-acarallexiviruses are not known (Mansouri and Ryšánek, 2021), plausibly most of the genera present here do not rely on winged insects for transmission. The exception is UlPolV1, which, as a polerovirus, is likely phloem limited and not only relies on aphid transmission in nature but isn't readily mechanically transmitted experimentally by the typical means (Miller, 1999). Accounting for other findings in oca, it has been found to be infected by the potexvirus papaya mosaic virus (Fletcher, 2001), by the nepovirus potato black ring virus (Jones and Kenten, 1981), the tepovirus potato virus T (Lizárraga et al., 2000), the tobamovirus Ullucus mild mottle virus (Fletcher, 2001) (tobamoviruses often being readily contact transmitted (Ilyas et al., 2022)), and the cheravirus arracacha virus B (Jones and Kenten, 1981) closely related to nepoviruses.

In potato production, seed potato production usually takes place in colder, upland areas where vector pressure is reduced (Kreuze et al., 2020a) and in potato systems in the Andes, viral degeneration is very low above 3,500 masl (Scheidegger, 1989). In 2001 a small number of oca lines were brought into Aoteoroa-New Zealand from various sites in the Andes: the paper announcing the lines mentions their altitudes may reach 3,500 masl (Martin et al., 2005). They were specifically subjected to elimination work that successfully produced oca plants which were free of infection from PapMV, AVB and ullucus mild mottle virus (UMMV) (Fletcher, 2001). These were noticeably absent from the samples subjected to HTS that were purchased from the internet. If the long-day adapted oca lines circulating in Europe that originate from these lines, that could go towards explaining why the European oca virome differs from that reported in the Andes in the past, i.e. that some reported viruses are absent. Their infection with novel viruses that are not likely to be aphid transmitted could be due to viruses slipping through the net in the early 2000s due to not being tested for via ELISA, or they could have entered the oca plants in Europe from another host. In the case of the potexvirus, the origin of potexviruses is not presently known, and they have a worldwide distribution (Fuentes et al., 2021), meaning that phylogeny based on the master sequences from a relatively small number of plants is unlikely to elucidate the origin of the infection in oca. Whether a virus speciates sympatrically or allopatrically influences congruence between geographic location and phylogenetic trees, for example (Pagan, 2018) and for these viruses, those assumptions have not yet been made. Generally, the vegetative persistence of all of these viruses (all were initially detected in extracts from leaf tissue grown from tubers) means they could have entered the plant line in its recent international ancestry or in the more distant past in the Andean region. Such an existence affects the evolution of the virus, as a virus which transmits vertically at high rates but not horizontally, such as in endornaviruses, is likely under very different selection pressures than a virus which frequently transmits horizontally (Sabanadzovic et al., 2016). An analysis of the sequences of

endornavirus coat proteins found no association between lifestyle and selection pressure on this region, however (Desbiez et al., 2011).

Several of these viruses were detected in samples from around Europe. This may mean one of two things. The first is that the oca plants are the source of this spread i.e. where oca is found, the viruses are being carried with them, long distance. The second option is that these findings may indicate that the virus is already circulating in Europe more broadly. or example, OCV1, which was detected not only in oca but in naturalised *Oxalis spp.* and in ornamentals. OCV1 was the only novel virus described during in this study which has a non-oca host. For the rest, the first hypothesis, that these are 'oca viruses' and circulating with oca, rather than indigenous viruses simply being detected in oca first, remains to be disproved. The original isolation host is not always the host which a virus has been associated with for the longest time, evolutionarily speaking (Stobbe et al., 2012).

Most troublingly from a risk assessment point of view, all samples from a commercial garden website were determined to be infected by one of even this small selection of viruses; some plants contained more than one. Even a consumer with some understanding of plant health may expect that material from an established grower is virus-free without needing to be told so, as one individual colourfully put it "They're not writing "E-coli free" on their lettuce but you just assume that it will be" (Marzano et al., 2021). In many cases, "plant blindness" means that buyers may not even be aware that plant health is something they should be concerned with; many Europeans are not aware of plant passports and plant health certification at all (Michi et al., 2023).

This chapter used HTS and qRT-PCR, as well as the available literature, to assess the virome of oca plants available for purchase through e-commerce in the UK. These tubers originated from several European countries and were found to be infected with many novel viruses, as well as ArMV and UlPolV1. The phylogeny and a putative genomic organisation for these viruses are presented here. These findings suggest that oca is host to many viruses which are not yet known to science, for which the risks are not yet known. Two viruses, OAV1 and OCV1, were found to have a high incidence in oca plants, and OCV1 was the only virus detected in opportunistic samples of UK flora. Future chapters will discuss the possible implications of these infections for UK plant health as well as from a virological standpoint.

Chapter 4. Samples from a Commercial Nursery in Ireland

4.1. Introduction

While chapter 3 covers plants sourced chiefly from informal internet trades, this chapter deals with material supplied by a commercial grower interested in biosecure trading. Growers within the wider horticultural sector cite a lack of ability to detect and remove pathogens as a concern even where good-will is assumed, and hard-to-detect pathogens present a medium-to-high risk of pest spread under the current system (Pollard and Marzano, 2023).

In the United Kingdom, growers show an interest in biosecure practices especially where they can provide added value to the consumer, but increased production costs deter growers, and increased cost of buying deters would-be ecologically minded buyers (Marzano et al., 2021, Dunn et al., 2020). An inability or unwillingness to implement good plant health practices may result if doing so causes a greater cost to buyers and/or growers (Jones, 2006). A number of growers identified that they would need more information before committing to any voluntary schemes (e.g. Plant Healthy in the UK), but did not outright reject the notion (Yeomans et al., 2021). A workshop bringing together scientists and growers found that while grower are keen to improve the quality of their products including by means of voluntary certification, they may be discouraged if certification schemes are perceived to be excessively bureaucratic or costly (Green et al. 2021).

Poor phytosanitary practice in nurseries can lead to accumulation of pathogens. Recently, the PHYTO-THREATS study found diverse *Phytopthora* isolates in irrigation water, drainage water and puddles from commercial nurseries after collaborating with commercial nurseries (Green et al., 2021). Close association of different plants within a nursery, alongside their vectors, provides unique niches for differentiation, hybridisation and host transfer (Brasier, 2005).

The utility of voluntary certification schemes is perceived by growers to be limited. Without legal oversight, such a system is only as strong as the 'weakest link' in the supply chain, and in some cases it would only take one non-compliant individual to cause an outbreak (Marzano et al., 2021, Pollard and Marzano, 2023). Regulation of any kind is only effective if it is adhered to (either by a choice based on the benefits or through regulatory authorities punishing non-compliance). Members of voluntary initiatives were found to comply with compulsory (legal) animal welfare standards better than non-members in Great Britain when inspected (Clark et al., 2016).

High throughput sequencing has been considered as a method to certify virus-free planting stock, particularly to investigate that material is free of both new and novel viruses (Villamor et al., 2019). HTS is moving increasingly towards validation for import and export certification

diagnostics as well, following inter-laboratory comparisons of HTS of ribo-depleted total RNA extracts and VALITEST (Gaafar et al., 2021, Trontin et al., 2021). A report on the topic was published in 2022 to facilitate adoption of HTS for diagnosis of plant pests, including viruses (Lebas et al., 2022). Further published studies have explored the validation of HTS for detection of viruses in specific hosts as well, such as grapevine, berries, and banana germplasm (Soltani et al., 2021, Villamor et al., 2022, Rong et al., 2023).

All of the material in this chapter was sent from a commercial grower keeping ARTCs in in the Republic of Ireland, who keeps the plants in post-entry quarantine. The plant material they have provided for this study comes from a number of sources, including other European nurseries. A portion were subjected to virus elimination work at the facility prior to this study. This differs from the material from the previous chapter (chapter 3), which was mostly sourced from informal or even illegal plants via the internet. When investigating the plants via HTS, a number of novel viruses and one quarantine virus were detected, along with the first complete genome of opium poppy mosaic associated virus.

4.2. Results

4.2.1. Morphology

Ulluco 'Colombia' tubers were very small, deep purple in colour. Mashua 'Blanca' has pale brown skin and creamy white flesh. The mashua plant from the commercial nursery in Ireland formed deep purple, almost black, tubers.

4.2.2. Oca

HTS datasets generated from extracts of the oca varieties 'Mellow Yellow', 'Apricot Delight' and 'Morten Smith Red' did not contain any viral contigs, which suggests that they are clean material. The 'Halford's Red' sample dataset contained a small (247bp) contig matching with 98% identity to beet ringspot virus.

4.2.3. Yacón

Potato yellowing virus Two of the four yacón plants were infected with potato yellowing virus (PYV). The HTS dataset derived from sequencing of RNA extracted from sample Yacon-NZ returned several contigs up to 8045bp in length, three of which Trinity identified as isoforms. The dataset derived from sample Yacon-Dimi contained two PYV contigs, of 2530bp and 3457bp respectively. Members of the genus *Ilarvirus* do not have any one segment this long; their entire genome is around 8kb and split between three segments of 3.5kb, 2.9kb and 2.3kb respectively (Bujarski et al., 2019). Investigation of the longer contigs suggested that they were chimeric reads attached together at the 3' UTRs. Ilarviruses have similar 3' UTRs shared between all three segments, with the last 61bp being identical across all three (Tzanetakis and Martin, 2005). This is a drawback of assembling from Illumina reads, which assembled chimeric contigs representing two to three different segments as one contig.

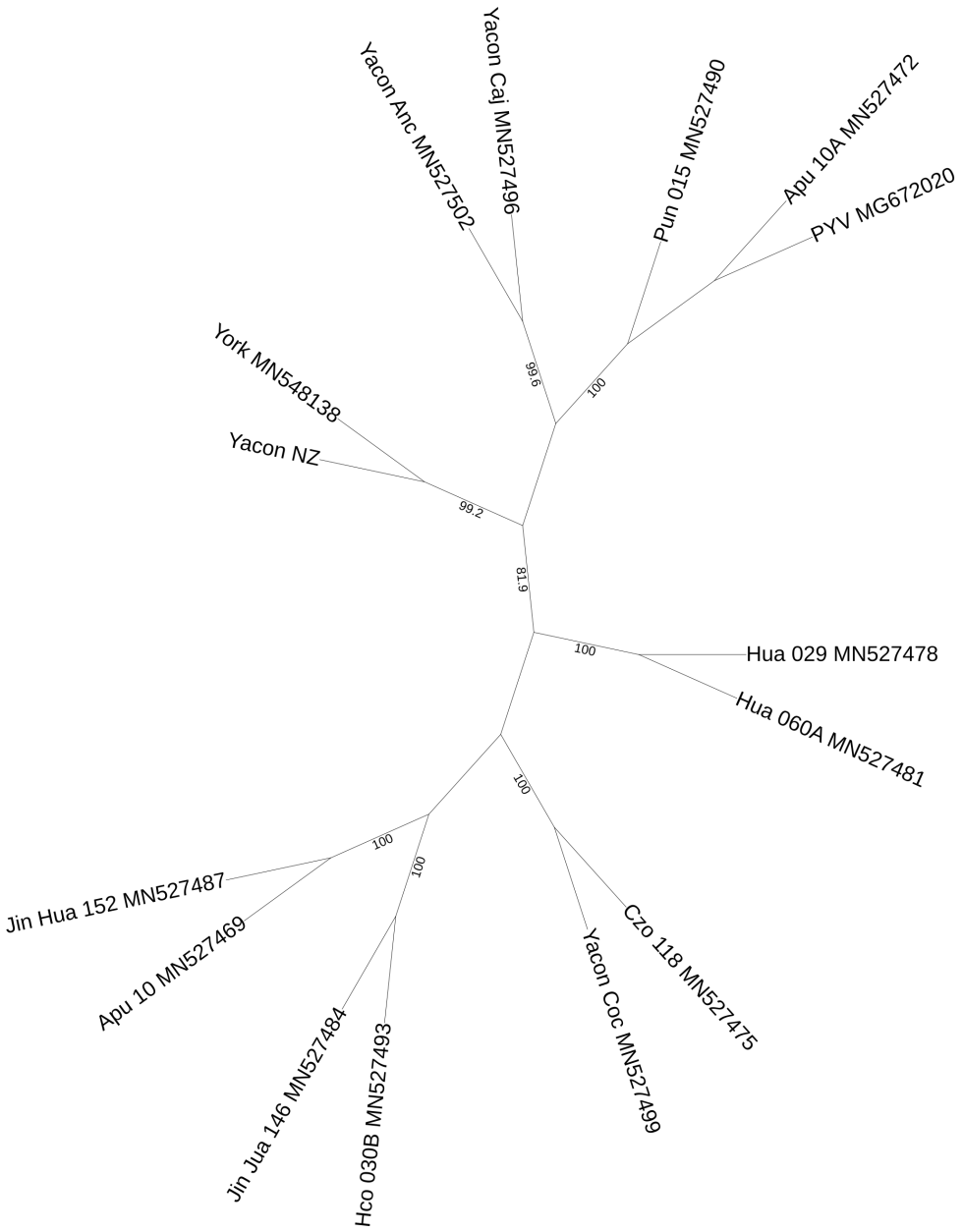


Figure 4.1 Unrooted phylogenetic tree of the RNA1 ORF of the PYV isolates, showing nodes with greater than 80% SH-alrt support. The two yacón NZ isolate clusters most closely with the previous European finding in York. Branch lengths do not indicate distance.

The other contigs present in the yacón HTS datasets were an 883bp a fragment of the coat protein (97.9% nt identity to accession MN544551, isolate PV-0706) from RNA extracted from sample Yacon-NZ and the complete sequence of RNA2 (95.4% nt identity to accession MN527503, isolate yacon-Anc) and RNA3 (97.1% nt identity to accession MN548139, isolate PVY-York at 93.92% coverage, 97.5% nt identity to accession MH937420, isolate PV-0706 at 93.7% coverage) sequenced from RNA extracted from sample Yacon-Dimi. Because the segments extracted from the same run were >99% identical (in some cases 100% identical), only one each was chosen for the phylogenetic trees.
In order to confirm the finding, RNA was extracted from new material sampled from the same plants. qRT-PCR and cnPCR were used to test both these new extracts and the original RNA extracted for HTS sequencing. The primers used were designed to detect the yacón isolate of PYV (subsection 2.3.1). The presence of PYV was not confirmed by qRT-PCR in any of the four extracts from PYV-positive yacón plants. Three of the extracts tested positive when amplified using conventional PCR, also using primers from Aimee Fowkes (Fera Science Ltd., unpublished data). These were both of the extracts subjected to HTS, as well as the second extract of Yacon-Dimi but not the second extract of Yacon-NZ. The other two yacón samples (Yacon-Morado and Yacon-Peru) were not found to be infected with PYV.

Caulimoviridae The datasets derived from extraction and sequencing of RNA from leaf tissue all four yacón plants contained various contigs matching *Caulimoviridae* member accessions with 82-96% identity, particularly yacon necrotic mottle (a badnavirus) and Dahlia mosaic virus strain DMV-D10, now known as Dahlia variabilis endogenous plant pararetroviral sequence (DvEPRS) (Eid and Pappu, 2014). These contigs ranged in length from 200 (the minimum contig length scanned via BLASTN using the Angua pipeline) to 800 base pairs. *Caulimoviridae* member sequences like these were also found to be common in yacón from the SRA archive (this thesis, section 6.2). One short contig matched another Cestrum yellow leaf curling virus with 90% identity.
One contig (234bp) matched Plasmopara viticola legion associated mononegaambi virus 8 (PvaMV-8) with 94% identity (94.8% coverage).

4.2.4. Ulluco

High throughput sequencing of RNA extracted from tubers in this study found Ullucus comovirus 1, broad bean wilt virus 2, Ullucus polerovirus 1, and papaya mosaic virus. A small 234bp matched sweet potato leaf speckling virus with 79% identity, potentially suggesting a novel polerovirus infection in at least one of the plants whose RNA was extracted and sequenced.

4.2.5. Mashua

Poleroviruses The mashua sample contained what is likely to be the genome of two different poleroviruses. In terms of their replicase, the first, tentatively named Tropaeolum polerovirus 1

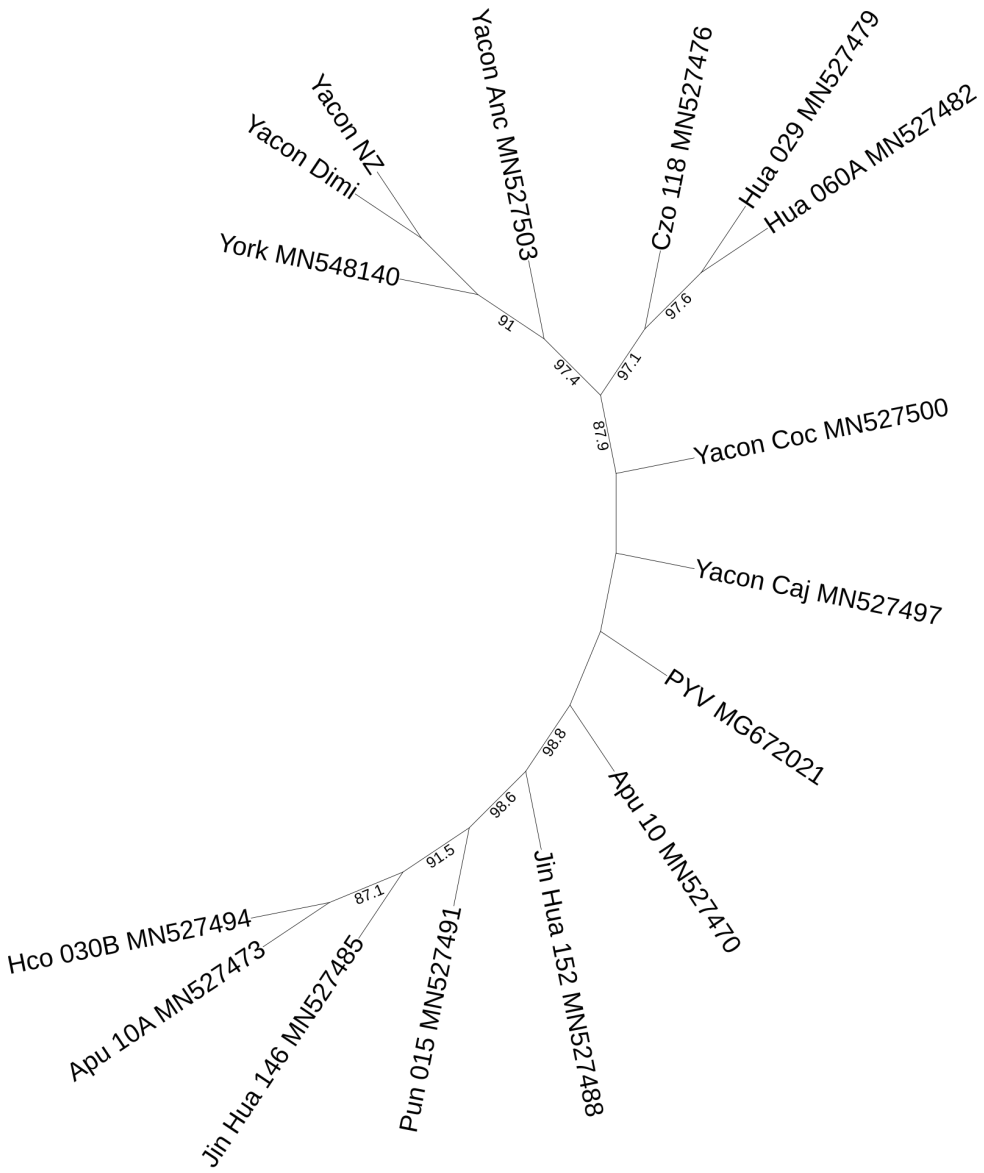


Figure 4.2 Unrooted phylogenetic tree of the RNA2 ORF of the PYV isolates, showing nodes with greater than 80% SH-alrt support. The two yacón isolates cluster most closely with the previous European finding in York. Branch lengths do not indicate distance.

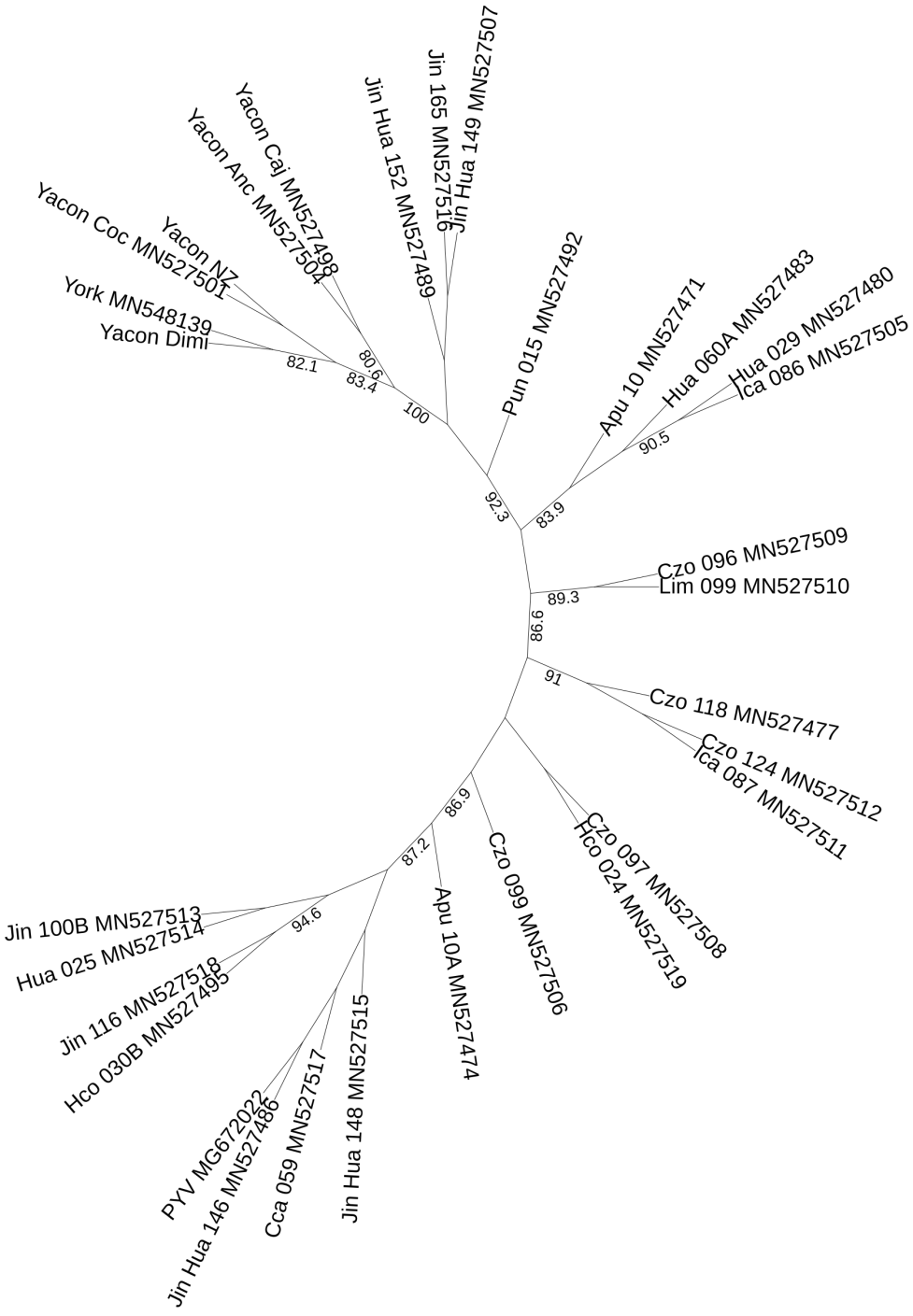


Figure 4.3 An unrooted phylogenetic tree showing the relationships between the nucleotide sequences of the CP genes of the PYV isolates. Bootstrap support is shown only above 80%. The sequences from the samples from Ireland cluster with the previous finding from York and with Yacon-Coc.

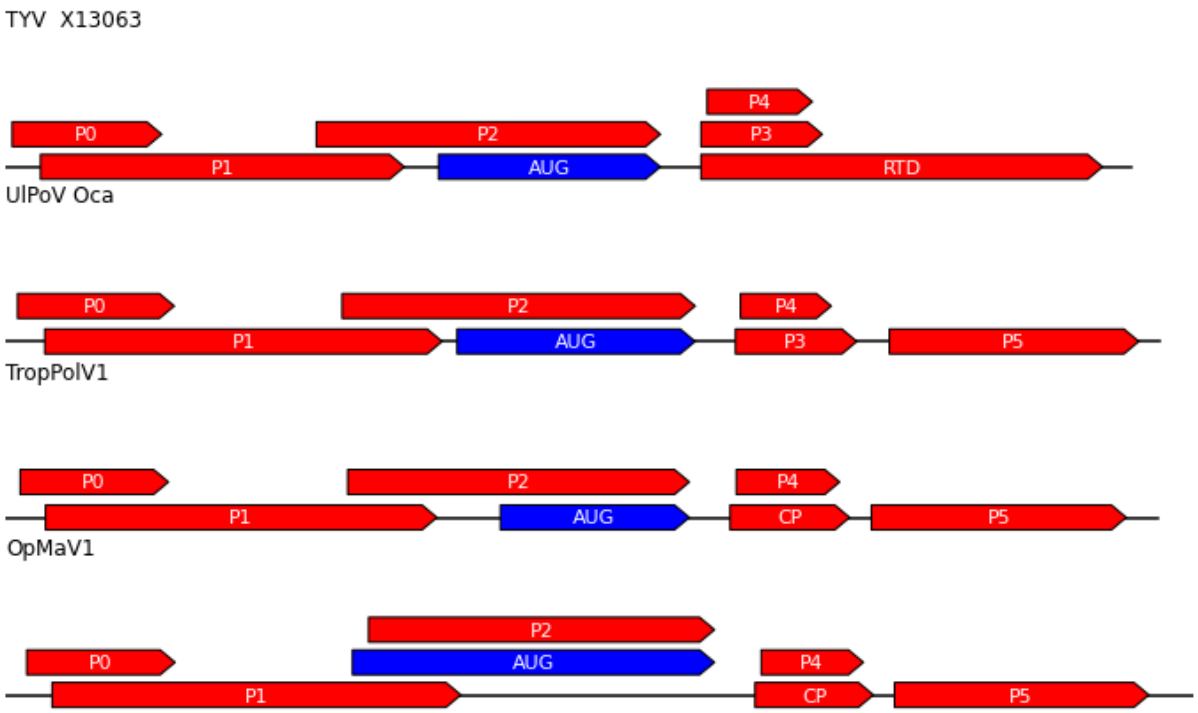


Figure 4.4 Diagram of the open reading frames of the exemplar accession of tulip yellows virus, the two polerovirus, opium poppy mosaic associated virus' complete genome and the oca isolate of Ullucus polerovirus 1. Note that the actual predicted frameshift ORF differs from the AUG start codon sequence in each instance. The tulip yellows virus annotations are from GenBank, excepting the AUG start codon, which was predicted with Geneious Prime.

(TropPolV1), is most similar in structure and predicted amino acid sequence to potato leafroll virus, with 70% amino acid identity in its ORF1 gene.

The CP/MP region sequence of the second polerovirus is 97% identical to that of opium poppy mosaic associated virus (OPMaV) isolated from *Papaver somniferum* in Aoteroroa-New Zealand alongside opium poppy mosaic virus, an umbravirus (Tang et al., 2016).

For both poleroviruses obtained by extracting and sequencing the RNA from the mashua sample, the second predicted ORF with an AUG start codon was disregarded. The predicted frameshifted fusion product was annotated as ORF2 and used for phylogenetic analysis. The slippery sequence was identified using KnotInFrame (Janssen and Giegerich, 2015) and alignment with other polerovirus P1-P2 regions. This region is putative as it was not tested for expression experimentally.

Querying the Serratus database with the frame 2 amino acid sequence of both species led to: no hit greater than 86% for the catalytic core of the replicase of opium poppy mosaic associated virus.

Analysis of the TropPolV1 sequence using Serratus PalmID against the SRA archive suggested that runs from a survey of potatoes in Kenya (SRR10448336), potato from Ireland (SRR11431599), potato from St. Petersburg, Russia (SRR8457041) human gut metagenome from Hangzhou, China (SRR7880367) and *Solanum chacoense* from Inner Mongolia, China (SRR10549512) had a >90% aa identity in the palmprint region; these were chosen for an initial comparison as they represented a wide geography and different hosts. Mapping of these reads to

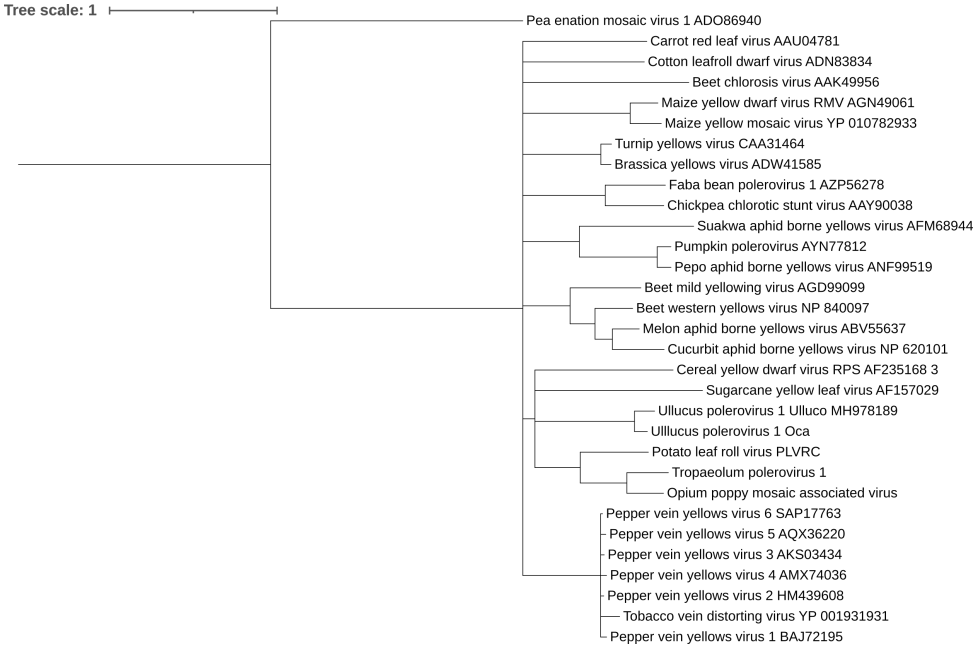


Figure 4.5 A phylogenetic tree of the novel polerovirus and opium poppy mosaic associated virus with other members of the genus, using the predicted replicase amino acid sequence.

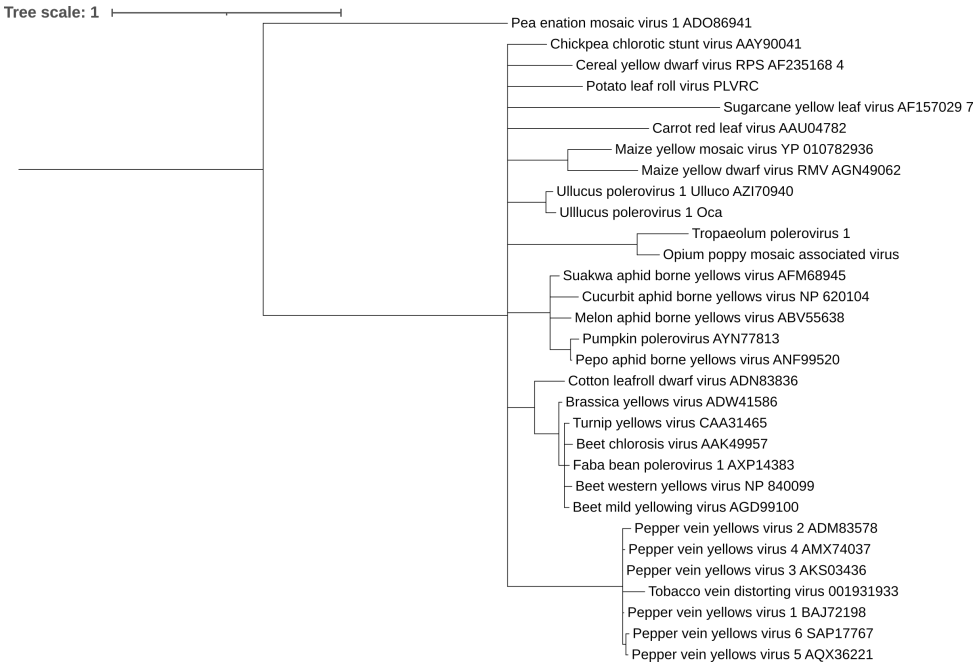


Figure 4.6 A phylogenetic tree of the novel polerovirus and opium poppy mosaic associated virus with other members of the genus, using the predicted coat protein amino acid sequence.

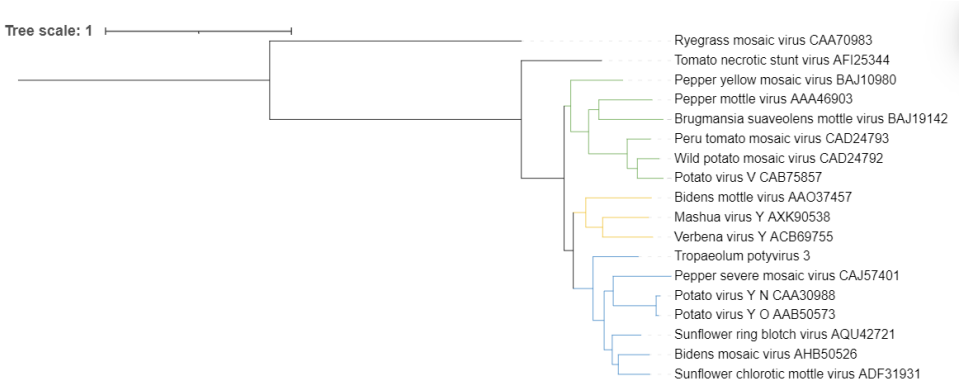


Figure 4.7 A phylogenetic tree Tropaeolum potyvirus 3 compared with other PVY subgroup potyviruses using the amino acid sequence of the entire predicted open reading frame. Individual clades are coloured yellow, blue and green.

the TropPolV1 genome did not result in a significant number of hits other than this individual region, and extraction of polerovirus reads from the assemblies suggested that the plants whose RNA was sequenced to form these SRA entries were all infected with potato leafroll virus and brassica yellows virus, neither especially closely related to TropPolV1. As this 90% identity reflected a between-species relationship for this virus, other 'close' (approximately 90%) hits were not extracted and assembled as well.

Potyviruses A number of contigs in the mashua-derived HTS dataset matched to potyvirus sequences, four of which likely originate from an infection with Mashua virus Y, with 87.5-96.9% nucleotide identity.
HTS of extracts of RNA from the mashua plant led to assembly of two sequences likely to belong to different novel potyviruses, one coding complete and another 1.2kb fragment.
The first sequence, taken to represent a novel virus named here as Tropaeolum potyvirus 3 (TropPV3) is likely to represent the entire genome. This sequence variously matched sunflower ring blotch virus (SuRBV), sunflower chlorotic mottle virus, Bidens mosaic virus and pepper severe mosaic virus, in the NIb and CP gene regions (up to 30-82% amino acid identity). These viruses belong to the potato virus Y subgroup of closely related potyviruses (Duarte et al., 2014, Quenouille et al., 2013). There are almost thirty members of this group, with PSMV likely being the most recent relative of PYV, diverging 8000 years ago (Fuentes et al., 2019).
Across the entire polyprotein, the novel potyvirus from mashua is most similar in amino acid sequence to that of SuBRV, at 71% amino acid identity, with which is also shared the most similar NIb protein, at 85.2% similarity with the type isolate. Figure 4.9 on Page 108 shows a summary of different members of the potato virus Y subgroup that are closely related to this virus, and their relationship in amino acid sequence of different putative proteins to this potyvirus from mashua. Notably, this sequence shares a coat protein that is around 80% similar to Bidens mosaic virus. A comparison between nucleotide sequence is also provided, as sequencing errors can lead to outsize differences in amino acid sequence for potyviruses (Adams et al., 2005b).

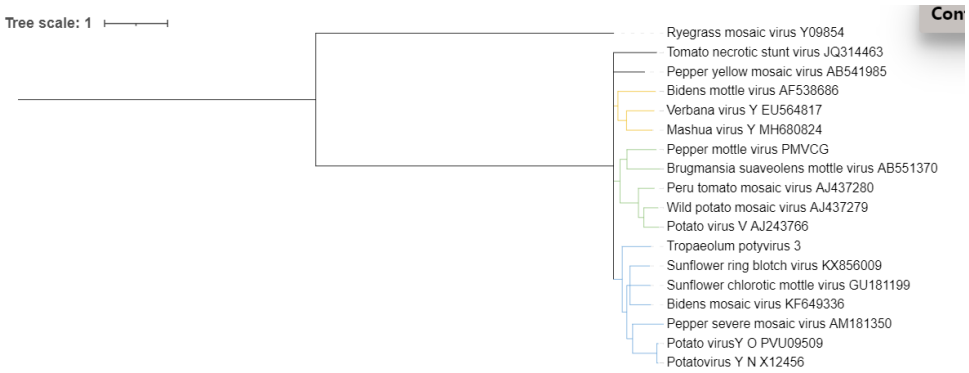


Figure 4.8 A phylogenetic tree Tropaeolum potyvirus 3 compared with other PVY subgroup potyviruses using the nucleotide sequence of the entire predicted open reading frame. Individual clades are coloured yellow, blue and green.

Figure 4.9 A figure showing identity in different peptide regions of the polyprotein between the novel potyvirus and other members of the PVY subgroup as calculated by SDT. The bottom set shows the same data with colours altered to prioritise demonstrating relative relatedness i.e. the closest relative per gene, and so the colour scales differ between genes.

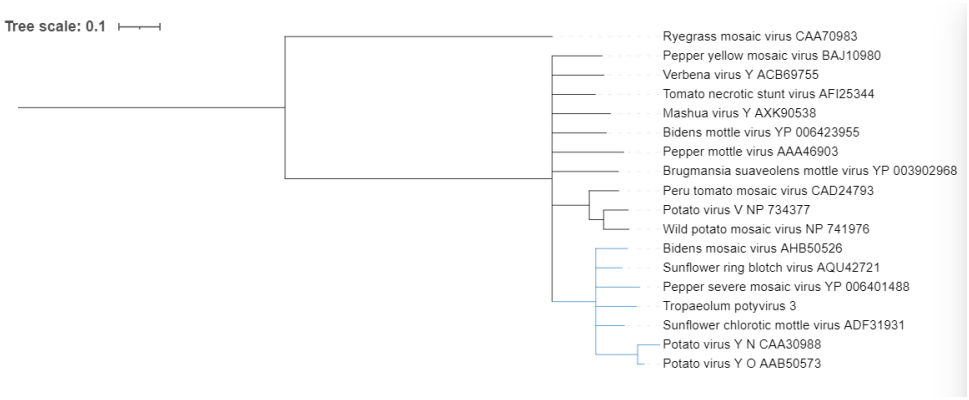


Figure 4.10 A phylogenetic tree Tropaeolum potyvirus 3 compared with other PVY subgroup potyviruses using the amino acid sequence of the predicted NIb mature product. The blue clade corresponds to the blue clade in other trees.

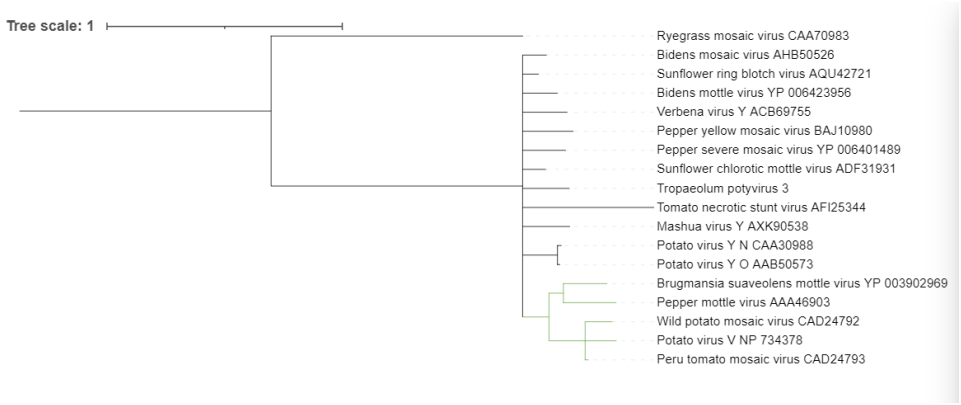


Figure 4.11 A phylogenetic tree Tropaeolum potyvirus 3 compared with other PVY subgroup potyviruses using the amino acid sequence of the predicted CP mature product. The green clade corresponds to the green clade in other trees.

Comparing the amino acid and nucleotide phylogenies of the entire polyprotein ORF, the two broadly concur with one another in terms of overall clades. The exception is that pepper yellow mosaic virus is not in the same (yellow) clade as Bidens mosaic virus in the nucleotide-based phylogenetic tree but sits within the overall potyvirus cluster. While individual sub-clusters differ slightly between the clades of the two trees, TropPV3 does not strongly cluster with any other member of its clade (blue, containing both compared strains of PVY) on any tree inferred. Comparison of the NIb amino acid sequence from TropPV3 and other potyviruses in the subgroup retains the blue (PVY) clade unaltered from either nucleotide or amino acid sequence of the entire ORF, though other clades are diminished and not strongly supported, such as the previous wild potato mosaic virus (green) clade which now contains only PVV, Peru tomato mosaic virus and wild potato mosaic virus. Comparison of the CP amino acid sequence preserved only the green (wild potato mosaic virus) clade and the grouping of the two PVY strains.

A comparison of the 3'UTR sequence did not produce any strongly supported clusters, other than situating the potyviruses away from the ryegrass mosaic virus outgroup, a rymovirus (data not shown).

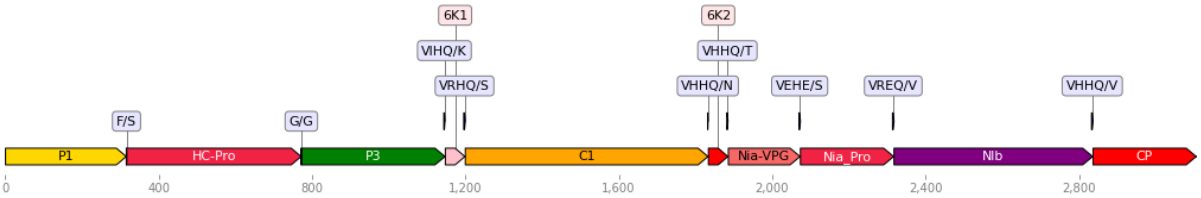


Figure 4.12 A diagram showing the putative cleavage sites and organisation of the predicted polyprotein from the novel potyvirus. Where a sequence leading up to the cleavage site is conserved in potyviruses, this is also annotated.

The putative genome organisation of TropPV3 is given here. Potyviruses, as with most other members of the *Potyviridae* family, produce ten mature polyprotein products from a polyprotein precursor (Adams et al., 2005b). The predicted cleavage sites of the novel potyvirus polyprotein are visualised in Figure 4.12. Alignment with other potyviruses suggests an F/S cleavage site for the P1 protein (S1), with a methionine four bases before the site, which is the same as experimentally confirmed for PVY (Adams et al., 2005a, Yang et al., 1998). The Hc-Pro (S2) cleavage site is YRVG/G. The sequence data suggest that both the NIa-Pro/NIb cleavage site and NIb/CP sites cleave between a glutamine and a valine; in potyviruses these two sites are both cleaved by the same protease (Goh and Hahn, 2021). This is supported in context for the NIa-Pro/NIb cleavage site by the VRE before the Q/V, predicted to be present in several potyviruses, and the tryptophan 37 amino acids before that that is highly conserved and the VGxH motif (Adams et al., 2005a). Although the site does not match the consensus found by Goh and Hahn in 2021, in rare cases valine can be the final acid in the motif, when it is between glutamic acid and valine, though not glutamine and valine (Goh and Hahn, 2021). Further confirmation of the sequence using Sanger sequencing would help determine if the predicted cleavage site is likely to be real or is an artefact of HTS. This would be the only potyvirus to cleave at Q/V.

A search on Serratus found that the predicted NIb protein has 92.9% identity with the corresponding region (catalytic core) of the reference (exemplar) accession of SuRBV (YP_009351881.1). This comparison supports the above finding that the NIb sequence of TropPV3 is most similar to that of SuRBV. The search also found that no SRA datasets contained no sequences of 90% or greater identity.

The other potyvirus sequenced via HTS from RNA extracted from the mashua sample was a 1221bp fragment of a potyviral polyprotein which, when translated in frame, has 79.8% identity to the P1 and HC-Pro region sequences of mashua virus Y. As the sequence recovered via HTS did not contain the replicase region, th sequence was unable to be compared with other RdRp catalytic cores using Serratus PalmID. Neither CP nor replicase protein products were represented in the recovered sequence, in their entirety, in order to perform a phylogenetic comparison.

Lastly, some sequences derived from the mashua HTS dataset matched 87-99% with mashua virus Y, representing coverage of the majority of the genome between them. The 1221bp sequence could be a divergent isolate of mashua virus Y infecting the same plant or could be a

novel virus; further work is required to obtain more of the genome of this virus and assign it to a species.

Betanucleorhabdovirus Two contigs from the same dataset, of 1495bp and 1037bp, had distant homology to accessions of Sonchus yellow net rhabdovirus when analysed with BLASTX. The 1495 bp contig matched the partial RdRp region via BLASTX with 66.9% identity in amino acid sequence and the 1037 bp contig matched the phosphoprotein at 60% identity, with a small portion of the N protein represented at the beginning of the contig.

4.3. Discussion

To begin, some of these data underline the importance of a 'human touch' in bioinformatics. The poleroviruses exemplify this: one of them incorrectly assigned to the genus *Umbravirus* by MEGAN due to an incorrectly labelled GenBank accession. Similarly, it is important to compare multiple regions of a draft genome for this reason. Taken as a whole, only 8% of the sequence representing the OPMaV genome matches to the known sequence of OPMaV, as the accession on GenBank is only a small portion of the genome and not present as a protein accession. Indeed, the replicase isn't represented at all, and so comparisons relying only on the RdRp region or BLASTX (which compares amino acid sequences) would have missed the relation and falsely attributed the sequence to an entirely novel virus.

4.3.1. Oca

The majority of the plants provided by the commercial supplier were determined to be virus-free. The exception, plants of variety 'Halford's Red', contained at least one plant that may ave been infected with beet ringspot virus (BRSV). BRSV is present in the European Union and not presently regulated there (Fowkes et al., 2021a). In addition, BRSV has been reported in *O. regnellii* in the united states (Jordan et al. 2019).
In total, 25 plants were tested for viral presence using RNA extraction and HTS. At least 20 of those plants were virus-free and of those infected, the only virus detected was not South American in origin. This suggests that, in comparison to the material from the internet (chapter 3) it is important to seek good quality sources of planting material when growing ARTCs, in order to avoid spreading viruses.

4.3.2. Yacón

All four of the yacón plants tested positive for DvEPRS. DvEPRS is understood to be endogenous, and was first discovered in *Dalia variabilis* (an aster) and was was first EVE found belonging to the genus *Caulimovirus* (Pahalawatta et al., 2008). It was subsequently found to be common in members of the family *Asteraceae* across several tribes (Maricle and Gillock, 2020). *Smallanthus sonchifolius* is a member of the *Millerieae* tribe within the *Asteraceae* family, but

this tribe was not sampled in the above study (Yin et al., 2023, Maricle and Gillock, 2020). This finding thus expands the number of tribes with a positive confirmation of DvEPRS.
Potato yellowing virus is qualified as an A1 pest by EPPO, meaning that it is not present in the region and is regulated (EPPO, 2023a). Since these findings come from post-entry quarantine, they do not represent a plant health risk, but they do indicate (along with the SRA results (section 6.2) and a paper studying the variation of PYV (Silvestre et al., 2020)) that PYV is reasonably common in yacón, including that circulating in Europe. It is partially for this reason that import of yacón plants into the UK is prohibited by DEFRA at present (Barker, 2022, DEFRA, 2024).
Yacon necrotic mottle virus has been associated with symptoms of necrosis, stunting and chlorosis in yacón before (Lee et al., 2015). The individual plants sequenced in this run did show signs of necrosis and chlorosis, but were both a) co-infected with other viruses and b) shipped as whole plants in packages before being sequenced. As a result, any association between the symptoms and individual viruses found is not possible to determine.
Plasmopara viticola, the isolation host of PvaMV-8, is a downy mildew oomycete (Chiapello et al., 2020). Presuming that the sequence derived from the yacón plants in this study is not contamination, it likely originates with an oomycete mildew infection on the yacón plants; white lesions were observed on the plant leaves, so this is entirely possible. *P. viticola* is typically a pathogen of grapevine and has a worldwide distribution (Taylor et al., 2019). PvaMV-8 was first discovered from sequencing of oomycete lesions on a Spanish grapevine (Chiapello et al., 2020). Thus, the virus sequenced here is likely infecting an oomycete on the yacón plants, and not the yacón plant itself. The finding of PvaMV-8 associated with these yacón samples is thus likely to be of minimal phytosanitary concern by itself, though the implied downy mildew infection may present a phytosanitary issue.

4.3.3. Ulluco

All viruses, except for the fragment of a possible novel polerovirus, were all previously detected in European ulluco (Fox et al., 2019). Ulluco is not presently permitted for import, due to a finding of potato black ringspot virus (PBRSV) associated with UK ulluco plantings (DEFRA, 2024). PBRSV was not detected in the five ulluco tubers or any other ARTC sample tested for the present study, nor were other A1 pathogens (EPPO, 2023a). However, some of the viruses detected in ulluco which are not quarantine are still of plant health concern. Broad bean wilt virus 2 (BBWV2) is an aphid-transmitted virus that causes economically important damage on agricultural and staple crops, such as broad bean (Ferrer et al., 2011). It is not clear whether the BBWV2 population infecting the ulluco plants originates in Europe, South America, or elsewhere. In South America, a complex of viruses in ulluco (Ullucus C comovirus, Ullucus mosaic potyvirus and papaya mosaic virus Ulluco strain) has been found to reduce yields by 38% (Lizárraga et al., 1999b).

4.3.4. Mashua

The mashua sample studied via HTS in this chapter was infected with several viruses in complex, some of which are likely to belong to novel species. Along with these putative novel viruses, this thesis reports what is likely to be the first full genome of OPMaV, the first finding of OPMaV in Europe, the first infection reported in *T. tuberosum*, and the first detection without a co-infection with the umbravirus, opium poppy mosaic virus.
OPMaV is incorrectly listed as an umbravirus on NCBI, but its associated paper lists it as sequence belonging to the luteovirid helper virus sequenced from an amplicon obtained via degenerate luteovirus and polerovirus primers (Tang et al., 2016).
Tang et al. (2016) used these generic primers to detect a polerovirus or luteovirus, probably OPMaV, associated with *Tropaeulum majus* (a common weed in the country) in the same location as the original *P. somniferum* plants (Tang et al., 2016). However, they did not associate the polerovirus to any symptoms by itself, as the scope of the paper was on the umbravirus detected within the plant and its mechanical transmission.
In the sequence of both poleroviruses (Tropaeolum potyvirus 1 and OPMaV) detected in the RNA extracted from these mashua samples, the intergenic regions between ORF2-3 were closer to 200nt than to 100nt, which is one differentiating factor between poleroviruses and luteoviruses (Domier et al., 2002). The Product 3 (P3) ORF of both genes ends in an amber (T/UAG) stop codon, which is suppressed in members of *Solemoviridae* to produce a readthrough protein (Sõmera et al., 2021).
Poleroviruses typically express P2 as a rare fused P1-P2 transcript at 1-30% frequency via a -1 frameshift; P2 contains the RdRp catalytic core portion of the replicase (Delfosse et al., 2021). In poleroviruses these two ORFs overlap by 400nt or more based on the presence of a 'shifty' sequence (Domier et al., 2002). In a number of cases the NCBI accessions are annotated with the AUG start-codon ORF2 rather than with a P1-P2 fusion protein whose P2 region begins partway through P1. In some accessions the predicted AUG-start P2 ORF does not overlap with P1, for example. Many of the NCBI accessions are annotated with the AUG-start ORF, such as accession MN689364-MN689394 (metagenomic assemblies from an unpublished paper on potato in Western Kenya), accession MK116549 (assembly from an unpublished paper on *Solanum phureja* in Colombia), and accession OR536956 (assembly from an unpublished paper on coriander in Chile).
In the literature, ORF2 is usually cited as overlapping with ORF1. Orfanidou et al. (2021) sequenced a number of Greek poleroviruses using overlapping PCR amplicons. Orfanidou et al. report using NCBI ORFFinder to detect the overlapping ORF1 and ORF2 (Orfanidou et al., 2021). A search using the same software found that the ORF corresponding to P2 (i.e. with an AUG start codon) did not overlap P1 in the isolates mentioned. The ORF discussed in the literature is thus likely to be the frameshifted protein product and is unrelated to the location of any AUG start codons, which are not expressed except as a frameshifted fusion product (Delfosse et al., 2021). This may have led to incorrect automated annotations on GenBank and elsewhere based on AUG-start ORFs that do or do not overlap with P1. This confusion is

apparent with the sequences of the two poleroviruses reported infecting mashua here: OPMaV1 has an overlapping P2 ORF and TropPol1 does not when accounting for the AUG start codon (Figure 4.4). Seeing the AUG start codon overlapping by several hundred nt, one may think that this is the true beginning of the ORF, though it is not actually expressed this way.
Phylogenetically, both poleroviruses (novel and additional genomic information of OPMaV) cluster together within the polerovirus group. In the replicase gene there is strong support for clustering the two new genomes with potato leafroll virus, though in the coat protein this clustering is not seen in the consensus or collapsed trees. This difference in clustering based on different gene products could indicate a recombinant origin of one or both viruses, or else potentially differing selection pressures on the two genes.
The assembly of PLRV and PYV from a human gut genome is interesting, but perhaps not altogether surprising. Plant RNA viruses have been found in human faecal samples before and have even been infectious (Zhang et al., 2005).
In terms of the putative TropPolV1, its replicase gene sequence clusters with viruses first isolated from potato and ulluco and with the oca isolate of Ullucus polerovirus 1. The ARTC and potato viruses form a clade that correlates with the ecological niche of the hosts. A previous study of poleroviruses found that the phylogeny of the genus correlates to the family of the host plant (Latourrette et al., 2021). Here the similarity is the niche of the hosts rather than their family (variously *Solanaceae, Oxalidaceae, Tropaeoleceae and Basselaceae*). Sequence representing *Opium poppy mosaic associated virus*, a species which was first isolated from poppy and whose members were detected also in *T. majus* from Aoteoroa-New Zealand (Tang et al., 2016), was also found here to cluster with poleroviruses first isolated from potato and ulluco. This clustering may suggest that *Opium poppy mosaic associated virus* is a species associated with *Tropaeolum sp.* whose members recently expanded their host range to opium poppy, rather than the other way around. The *Tropaeoleaceae* family originates in the Americas (Bayer and Appel, 2003). *T. majus* was a popular ornamental in of the Incas and was used for its peppery flavour and medicinal qualities in what is now Mexico (National Research Council, 1989, Hernándo Bermejo and León, 1994). Later *T. majus* was spread by the Spanish to Europe, and from Europe to other countries as the ornamental garden nasturtium, in some cases escaping into the wild as a weed (Tang et al., 2016, Hernándo Bermejo and León, 1994). Based on the available information it is highly likely that the ancestor of the poleroviruses isolated from European ARTCs in this study and in 2019 (Fox et al., 2019) is of South American origin. This ancestor or individual species then spread with the plants.
It is less parsimonious (though not impossible) that these botanically diverse plants would become infected with European and Australasian viruses that just so happen to be closely related, on entirely different continents. OPMaV might have spread to opium poppy via the naturalised *T. majus* weeds, suggesting an earlier introduction, or perhaps from mashua plantings in the more recent past. A virological survey of *T. majus*, mashua and poppy in different countries (e.g. Aoteoroa-New Zealand, Europe, Mexico and South America) may assist with elucidating this relationship.

The ICTV species demarcation threshold for poleroviruses is >10% amino acid difference in any one gene product (Sõmera et al., 2021). Based on these findings, the species name *Polerovirus tropaei* (Tropaeolum polerovirus 1, TropPolV1) is suggested for the species for which a complete genome sequence is likely to be represented here. The other sequence found is likely the first full genome representing the species *Opium poppy mosaic associated virus*, which has no latinised binomial at present and is not ratified.

Two novel potyviruses were also detected infecting the mashua plants, with one sequence representing the apparently complete genome of a novel potyvirus. Of the two novel potyviruses detected in the mashua sample, either could be one or both of Tropaeolum tuberosum potyvirus -1 or -2, which were frequently found in complex. Tropaeolum mosaic virus (TropMV) was reported at the time of its first description to be very common on mashua plants. Mashua virus Y was likewise found in both a historic isolate and an internet purchase in 2018 (Adams et al., 2018c). Further detection, through HTS, of MVY infecting during the present study may be an indicator that these two viruses are the same. However, this would only be the parsimonious hypothesis, not necessarily the correct one. TropPV3 from this paper could also be TropMV, though it would appear to be less widespread in mashua plants than TropMV is reported to be. Biological and molecular characterisation of MVY or TropPV3 to confirm similar molecular and serological properties to the two previous *Tropaeolum* potyviruses would be required or, ideally, sequencing of historical material labelled as TropMV.

Neither Tropaeolum potyvirus -1 or -2 were found to be related serologically to the '17 other potyviruses' (Brunt, 1990). However, antisera do not always distinguish closely related PVY subgroup species members, as is the case for SuRBV and a related virus, sunflower chlorotic mottle virus (Cabrera Mederos et al., 2017), isolates of which in turn did not react to PVY antisera (Dujovny et al., 1998). Tissue infected with SuRBV did cross-react with Bidens mosaic virus (Dujovny et al., 1998). TropMV-infected tissue did not cross-react to PVY-C (common strain) antisera. Mashua virus Y was found to be somewhat related, based on sequence information, to verbena virus Y, another PVY subgroup virus (Adams et al., 2018c).

Though seven*Potyvirus* species (*Tropaelum mosaic potyvirus*, an unnamed potyvirus causing mosaic symptoms, *Tropaeolum tuberosum potyvirus 1* and T*ropaeolum potyvirus 2* (chlorotic symptoms) (Brunt, 1990), *Mashua virus Y* and the two putative novel species described from this thesis) have been described in mashua, lack of cross-comparison between different findings means these could represent as few as three species. Lack of characterisation of novel species or means of associating previously described species to sequence are the reason for recent efforts to sequence historical samples, but some original isolates may have been lost (Jones et al., 2021).

The sequence here named as TropPV3 could be recombinant, or the differences in relationship between TropMV3 and other potyviruses based on the sequence of different genes could be due to the varying rates of evolution between genomic regions for members of the genus *Potyvirus*. For example, the P1 protein is the least conserved, and in TropPV3, the P1 region is indeed most divergent from sequences from other species (Adams et al., 2005b). Even within gene products conservation can differ, with the CP 'core' differing from the entire CP gene in its level of

conservation. Recombination is likely to be common in potyviruses. Recombination events have been reported extensively in potato virus Y in the CP and 3' UTR regions (Revers et al., 1996) and in members of the *Potyviridae* family more broadly in the highly variable P1 gene (Valli et al., 2007).

The NIb and CP regions of the TropPV3 sequence show greater than 80% identity to these regions of other members of the potato virus Y group. Other gene regions of TropPV3 are highly divergent. Notably, the TropPV3 NIb gene sequence is similar to that of other potyviruses known from Argentina and Brazil, such as pepper severe mosaic virus, sunflower chlorotic mottle virus, SuRBV, and chiefly Bidens mosaic virus (Bejerman et al., 2010, Dujovny et al., 1998, Cabrera Mederos et al., 2017, Feldman and Gracia, 1977). Bidens mosaic virus was detected via partial Nib and CP nucleotide sequence in arracacha in Brazil in 2017, with 95-98% identity in these gene regions to the *Bidens sp.* isolate (Orílio et al., 2017). TropPV3 plausibly originates in South America, probably entering Europe with imported mashua plants if so.

Potyvirus taxonomy is more complex than some other viruses (Bos, 1992). This is in part because, as mentioned, the taxonomic information to be gained varies per gene, with different genes having different levels of conservation (Adams et al., 2005b). In particular, the potato virus Y subgroup contains a number of viruses all closely related to *Potato virus Y*, some of which were considered strains of *Potato virus Y* in the past. *Bidens mosaic virus* and *Sunflower chlorotic mottle virus* have previously been considered strains of *Potato virus Y* but are now their own species (Inoue-Nagata et al., 2006, Sanches et al., 2014, Adams et al., 2005b, Cabrera Mederos et al., 2017). This is the same for *Pepper mottle virus, Pepper yellow mosaic virus,* and *Potato virus V* (Spetz et al., 2003). Coat protein sequence tends to distinguish between species but in some closely related species, amino acid identity may be above the threshold which would determine the isolates as members of the same species (Adams et al., 2005b). Population genetics studies may help to better differentiate potyvirus species/strain members than sequence alone by distinguishing separately evolving populations (Gibbs et al., 2020). Further work on sequencing of mashua plants may help to uncover more sequences belonging to this putative species and aid in such a study.

The ICTV species demarcation threshold for the *Potyviridae* family states that an amino acid identity of less than 82% in ORF products may be used to delineate one species from another (Inoue-Nagata et al., 2022). In the case of TropPV3, the majority of the predicted mature protein products are beneath this threshold when compared to other potyviruses. Based on this, TropPV3 is suggested to be a member of a novel potyvirus species *Potyvirus tropaei* (Tropaeolum potyvirus 3, TropPV3). The other potyviral sequence recovered may represent a member of a new species, but further characterisation would be required to be certain of this.

For the sequence likely belonging to a member ofthe genus *Betabucleorhabovirus*, both regions represented by the HTS-derived sequences have less than 75% identity with Sonchus yellow net rhabdovirus. It is possible that these sequences represent a novel virus in the genus *Betanucleorhabdovirus*. However, the combined contigs represent only 2532 base pairs of coverage of the putative genome. The typical genome of a betanucleorhabdovirus is 13-15kB

long (Walker et al., 2022). Thus only 17-19% of the genome has been recovered here, and for a full identity comparison the entire genome is needed. Further sequencing to obtain the full genome sequence of this putative novel betanucleorhabdovirus will be required (particularly the full replicase) to confirm that this is a novel virus within the genus.

4.3.5. Conclusion

The work presented here represents a collaboration between a nursery with an interest in good phytosanitary practice and a laboratory which can provide non-targeted diagnostic support through HTS. Such detections and diagnoses might be carried forward to relevant NPPOs, thus saving some degree of time and expense when seeking formal certification. It may also be used for compliance with voluntary certification schemes which may not require an accredited laboratory or NPPO to issue a certificate in a formal capacity.

At present, legal certification of propagative materials can only be performed by specific NPPOs; in the UK this is Department for Environment, Food and Rural Affairs (DEFRA), and in Ireland this is Department of Agriculture, Food and the Marine (*An Roinn Talmhaíochta, Bia agus Mara*) (https://www.ippc.int/en/countries/all/list-countries/). Under EPPO guidance this certification system must follow specific, validated protocols by accredited laboratories (Trontin et al., 2021). In the United Kingdom, a 2023 publication has announced the intention to introduce the concept of 'Trusted Traders', where importers are authorised to certify their own materials subject to audits (DEFRA et al., 2023). Allowing the private sector to self-regulate in this way is somewhat controversial, due to what may be seen as competing interests between financial gain and truthful reporting of infection (Jongeneel and Herzfeld, 2012). However, nursery proprietors are broadly incentivised to produce healthy material due to what is largely a trust-based system of exchange. In the UK, most contracts are verbal only, by so-called 'handshake' and rely on a complex system of trust and obligation (Pollard and Marzano, 2023). On the other hand, high throughputs, short deadlines and an incentive to reduce costs (both for themselves and to keep a competitive pricing for consumers) can inversely incentivise lax practice in horticultural companies, such as purchasing from suppliers without pre-vetting their standards (Pollard and Marzano, 2023).

Viral infections are the most common pathogenic organisms of ornamentals and are often asymptomatic, but outbreaks on symptomatic hosts may cause economic damage (Mahmoodi Safa, 2016). In some cases businesses which purchase infected material may be able to seek recompense from those who sold it, though not all suppliers honour this system (Pollard and Marzano, 2023). In balance, then, it is generally desirable for growers of propagative material to maintain good health of their stock, but real-world concerns may make this an economic trade-off.

The use of HTS to sequence viruses infecting propagative material can thus act as a pre-screening. For example, the oca and yacón lines deemed to be virus free through this study may be bulked up in post-entry quarantine while resources are not devoted to bulking up plants diagnosed with PYV. The infection with novel viruses whose biological importance is not known

makes such decision-making more difficult but, in particular, plants diagnosed with an infection with known quarantine viruses such as PYV may be discarded in this manner. At present, diagnostic information obtained via HTS can be used to pre-screen material for certification from the relevant NPPO, avoiding unnecessary and potentially expensive formal tests on material which is already known to be infected. This approach will also allow a grower begin to pre-bulking material which is likely to be clean before release, saving time and resources and reducing turnaround times (Maree et al., 2018). In future HTS may form the basis for post-entry diagnosistics (Lebas et al., 2022) though work would be required still to distinguish integrated forms from actively replicating ones in the case of detections of DNA viruses (Brait et al., 2024).

Chapter 5. Host Range of Nepoviruses

5.1. Introduction

Historically, the study of viruses has favoured those which cause disease of crops and, crucially, those which are easily experimentally manipulated, e.g. through mechanical transmission and serial passaging (Roossinck, 2019). The ongoing revolution in virology presented by the sheer quantity of data that HTS can provide is changing this bias in many ways. Viruses new and known are being found in new hosts and countries due to HTS (e.g. (Fox et al., 2022, Sidharthan et al., 2022b, Blouin et al., 2023)). This reflects an overall divergence between "virology", which became increasingly molecular with time in the post-sequencing era, and "virus ecology", though the two have dovetailed once more with the scale and context HTS can provide (Maclot et al., 2020).

Such sequence-only findings as may derive from HTS make distinguishing between pathogenic and merely commensal viruses difficult (Adams et al., 2018d). Reports of novel viruses in hosts which may have been overlooked in the past often have relatively little biological information available, which has been referred to as the 'biological desert' (Hou et al., 2020). For novel tomato viruses discovered via HTS, only around 30% of the papers that were collated and analysed by Rivarez et al. in 2021 conducted indicator host range studies (Rivarez et al., 2021). Viruses are also increasingly discovered within metagenomic or environmental samples without an obvious host, meaning that some viruses lack biological information entirely (Simmonds et al., 2017). As plant pathology advances, plants and their interactions with other organisms, which may be pathogenic for some or all of their life cycle, is increasingly understood to have complex environmental, biological and human factors which need multidisciplinary study (Jeger et al., 2021).

Biological data concerning a virus are necessary for epidemiology and have phytosanitary implications - i.e. when is it appropriate to restrict trade based on a virus, given that it could be a) harmless and/or b) already present, unobserved, in the importing country (Macdiarmid et al., 2013, Olmos et al., 2018)? If a novel virus is found, characterisation takes time and resources, often more of both in the case of highly divergent viruses which do not resemble known viruses closely, for technical reasons (Fontdevila Pareta et al., 2023). A certain, minimal level of knowledge of a virus is required to determine when study of it is best 'shelved' in favour of higher-priority, riskier pests. However, any biological information which can be determined quickly and efficiently is worth obtaining; small scale epidemiological surveys of the original isolation field or batch to determine incidence are one such example, as well as primer design (Hou et al., 2020).

Determining the biology of a plant virus may also require obtaining a working, single-isolate population of the virus within experimental plants. Mechanical inoculation is a valuable tool for studying plant virus biology for this reason. It serves to quantify virus titre via local lesions, to allow study of early stages of infection, to allow propagation of virions for purification, separation of mixed infections, tests of infectivity and determination of a preliminary host range for targeted surveys (Hull, 2001, Steere, 1959, Ranabhat et al., 2022, Massart et al., 2017). In other words, the ability to transmit a virus from the isolation host to various other experimental or crop hosts is an important part of characterising a novel virus.

Uses of mechanical inoculation Mechanical inoculation precedes a number of other biological methods. The first use is isolation of a single virus population, usually a species or strain. This can be achieved through two main means, biologically speaking. Certain herbaceous indicator hosts, such as *C. quinoa*, produce characteristic local lesions when inoculated with many viruses due to the hypersensitive response (Zhang et al., 2012). These local lesions may be used for quantitfying titre (lesions counts) (Kleczkowski, 1949) or can be sampled from in order to separate out mixed infections through single-lesion isolation (single-legion transfer) (Steere, 1959). Usage of hosts with differing susceptibilities to viruses is another means of separating out mixed infections, and requires knowledge of the host range of the virus in question. If the host range of the virus or viruses is very broad, however, this latter method is not always available (Fox, 2020).

Isolation of a single strain or species serves several purposes and is arguably 'step one' for further characterisation. For example the virus may be re-inoculated onto a variety of the original host for symptom study, or onto other species to study the host range. Such isolation is also often utilised as part of the modified version of Koch's postulates commonly followed today in order to demonstrate back-inoculation, proposed in a concise form by Bos in 1981 (Bos, 1981). This will be discussed later in this thesis (see chapter 7) and is thus only lightly touched upon here. Where a suitable host is available, mechanical inoculation onto a host can indicate that the virus detected within the plant is both viable and infectious (Ranabhat et al., 2022). Methods such as PCR and ELISA are considered indirect and cannot distinguish viable from non-viable pathogens. Mechanical transmission, along with other horizontal transmission methods, may be attempted in order to help distinguish episomal from integrated forms of plant pararetroviruses (e.g. (Lockhart et al., 2000)). Failure to transmit may simply mean that the transmission methods failed, rather than that it is incapable of transmitting horizontally at all. That is to say, a negative result does not indicate a non-transmissibile form, but a positive result likely indicates an activated form.

Mechanical inoculation may aid in diagnosis. 'Indicator' hosts are those which produce characteristic symptoms when infected with a particular virus isolate and aid in diagnosis (Hull, 2001). In some cases symptom expression may be quantified (Ranabhat et al., 2022). At present diagnosis of viral infection is mostly serological or nucleic acid based (Trontin et al., 2021), though in the case of grapevine biological methods are still routinely in use as viruses that are known to infect grapevines can be highly variable at the molecular level (Velasco and Padilla,

2021). As early as the 1980s, host range was longer considered required or perhaps even reliable for virus identification (Hamilton et al., 1981). Lack of standardisation in host range testing has been a limitation on its usefulness as well (Brunt, 1990), along with as differences in host phenotype even within the same cultivar between laboratories (Van der Want et al., 1975).
Purification or propagation hosts are another type of host with use to the virologist. These hosts are those which readily accumulate virions inside their cells which are amenable to extraction in high concentrations for further study (Lane, 1986). Maintaining competent virus within hosts through serial passage or, less commonly, a perennial host, is a means to keep a source of inoculum on hand (Adkins and Rosskopf, 2002) (though see chapter 7 for some dangers of this when it comes to highly plastic viral genomes).
The last use to mention here is for investigating potential epidemiology. Though mechanical inoculation is not quite the same as the conditions under which a virus may transmit in nature (e.g. introducing much larger quantities of virions than an aphid in the case of PVY (Dupuis et al., 2019)), it does provide information on whether a virus may systemically infect a plant should it enter its cells through wounding. Such studies might help identify economically important secondary hosts of a novel virus, for example, and inform targeted field surveys for risk assessment (Massart et al., 2017).

Herbaceous host considerations A number of properties make a plant useful for virological study. These differ via use case as well.

- **Virus range**: An indicator should have a good chance of being infected by a virus when inoculated with it. This is the case for the likes of *Nicotiana benthamiana*, laboratory strains of which have a defective RDR1 gene involved in viral defence (Todesco and De Felippes, 2016). This needn't be all viruses, as an indicator's immunity to a given virus can allow differential host diagnosis and isolation of viruses from mixed infections (Fox, 2020).

- **Husbandry**: Indicator plants should be easy and reliable to grow. Some species of plants, especially weeds, exhibit dormancy only broken by scarification or vernalisation, and heterogeneity in germination timing (Penfield, 2017). An ideal host plant should germinate and reach a useful size quickly (Christie and Crawford, 1978) and uniformly.

- **Homogeneity**: Indicator plants should be true to type and homogenous in phenotype. One member of the species should ideally show the same reaction when challenged with a virus as another member of the species (Van der Want et al., 1975). This is especially important when using indicator hosts to demonstrate causation, as inconsistent symptom expression within a host population is a confounding factor in doing so (Fox, 2020).

- **Reactivity**: Production of large, countable local lesions and strong, characteristic symptoms is desirable for indicator hosts (Van der Want et al., 1975). A quick development of symptoms once inoculated is convenient and useful (Van Dijk et al., 1987).

Conversely, a maintenance host should not be overly reactive in the sense that it dies quickly. A plant that survives infection for an extended period of time is useful for maintaining inoculum when storage of infected tissue reduces infectivity (Adkins and Rosskopf, 2002).

- **Morphology**: Growth habit and leaf shape affect a plant's amenability to inoculation. Large, round, smooth leaves are physically easier to sap inoculate by rubbing than small, narrow or wrinkly ones (Van der Want et al., 1975).

The inoculation methods are covered in detail in section 2.6. Briefly, tissue found by real-time PCR to be infected with the novel virus candidates, ONV1, OVX and OAV1 was inoculated onto a suite of test plants by grinding with celite and applying using a gloved finger. In particular, ONV1 was tested on 140 C. sativa plants. 110 of these took place after a binomial distribution test determined that this was the number of plants needed to test the hypothesis that the single observation of an infected C. sativa plant was incongruent with the other 29 negatives. The oca isolate of Arabis mosaic virus was successfully transmitted in several cases, confirmed by PCR. For the host range review, the literature was consulted for each virus. Paper archive databases (SCOPUS, Google Scholar and Web of Science), CABI Compendium (CABI, 2024), and the EPPO Global Database (EPPO, 2023c) were searched using the species and virus names (including synonyms) of each ratified species in the genus *Nepovirus* (as of 2023). In the EPPO and CABI databases, the cited literature was also investigated. Citations within individual papers were investigated to form, as completely as possible, a picture of the available literature on the host range of these nepoviruses.
In this chapter, difficulties with biological characterisation of a novel virus from oca are used as a jumping-off point to discuss the role of biological characterisation when investigating viruses discovered via HTS, particularly with no associated symptoms on the host plant. A review of the literature is presented combining >125 sources to give an overview of the host range of 41 nepoviruses.

5.2. Results

5.2.1. Sap Inoculation

Out of the initial suite of plants tested, the only plant to test positive for ONV1 after sap inoculation was a single *C. sativa* plant. Based on this, a beta binomial distribution test was performed to determine the number of plants which would be needed to suggest that the initial finding was atypical statistically. 108 plants would be needed to determine this with a high confidence, or else disprove the hypothesis. Of the eleven sets of inoculations, none of the 110 plants tested positive, suggesting that the first finding was statistically atypical and represented a different set of criteria than the other ten samplings in some way.
A small number of plants were inoculated with positive samples of OAV1 and OVX (DK1). No allexivirus samples were positive with the respective viruses when tested with qRT-PCR primers, nor were any potexvirus-inoculated samples.

However, some plants inoculated with OVX from DK1 showed flecking and chlorotic spots on their leaves. One *C. quinoa* plant developed mottling after three days.
This was because, due to difficulties designing primers, the OVX+ plants also had Arabis mosaic virus and this had not been detected at the time of inoculation. Using the new ArMV primers on the sapped material, it was determined that the symptomatic cucumber plants had been inoculated with ArMV and not OVX.
Of the plants inoculated with ArMV, 4/6 *C. quinoa* plants tested positive, 4/5 *C. sativa* plants tested positive (and of those inoculated from an infected *C. sativa* plant, 5/6 tested positive for a total of 9/11), 4/4 *N. benthamiana* plants tested positive, and 0/1 *N. tabacum* plants tested positive. None of the *N. tabacum* plants which were not inoculated with virus tested positive for any virus. Positive controls of *C. sativa* inoculated with CMV, *C. quinoa* inoculated with ArMV, tomato spotted wilt virus inoculated onto *N. rustica* and *Impatiens sp*. and BBWV1 inoculated onto *C. quinoa* caused the inoculated plants to develop symptoms characteristic of the respective virus.
None of the material from tubers frozen in 2019 transmitted successfully onto any indicator plants. These were samples YfP1, YfP2, NP2, HHPP, FG and LO.

5.2.2. Literature Search of the Host Range of Nepoviruses

In total, the literature was consulted for 41 nepoviruses and 758 plant species, totalling 2133 host-virus combinations recorded.
In instances where two sources conflict, such as viruses of *H. annuus* in (Wingard, 1928) and (Hibben and Bozarth, 1972), they are marked with an asterisk indicating the opposite result was found at least once. A similar case occurs with the infectivity of TSRV in *Beta vulgaris*: two strains were found on separate occasions to infect *B. vulgaris* (Komuro and Iwaki, 1968, Price, 1940) but the ash strain was not found infective in *B. vulgaris*, confirmed by back-inoculation (Hibben and Bozarth, 1972). In the interest of collating the data, species have been collapsed together, i.e. if one strain infects the plant but not another, the species is still considered to infect that host.
Of the nepoviruses reported in this study, 27 of the 41 viruses were inoculated onto (or naturally found on) at least 20 species. This includes instances where an inoculation was attempted but reported unsuccessful. Only fifteen were tested or detected on forty or more plant species.
As seen in Figure 5.1, for each subgroup, one to three viruses dominate the number of plant species tested. These are TRSV (subgroup A), tomato black ring virus (TBRV) (subgroup B) and tomato ringspot virus (ToRSV) (subgroup C). A similar skew can be seen in the proportions of viruses tested on plant when divided by family: of 2133 observations, 424 were host-virus combinations where the plant challenged was a member of the *Solanaceae* family. In total 99 host families were tested, meaning that if the families were tested evenly, one would expect 21.5 tests per family. In other words, *Solanaceae* is represented almost 20 times more than would be expected if the distribution were even.

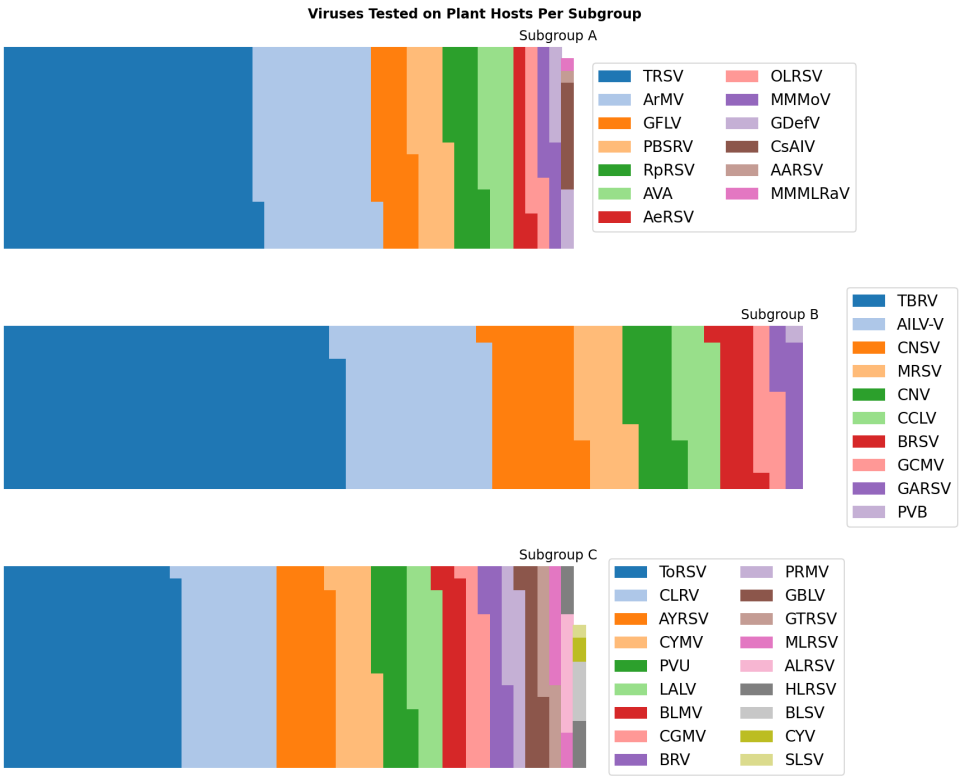


Figure 5.1 A waffle chart of host-virus combinations divided per subgroup. The area is proportionate to the number of host-virus combinations recorded, successful or unsuccessful, including natural infections. The colours are re-used between subgroups, i.e. blue represents TRSV in subgroup A but TBRV in subgroup B, this indicating the relative first, second, third most tested virus for each subgroup. In each chart 1 block = 1 virus-host pair.

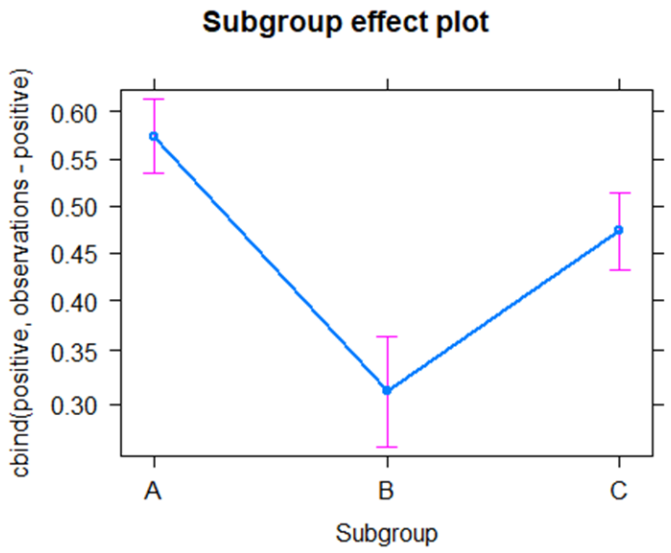


Figure 5.2 A chart showing the possible differences in host range between subgroups, as suggested by the data derived from the literature search. This figure and the underlying R script were provided by Roy Macarthur of Fera Science Ltd.

The five nepoviruses overall which are most well-studied in the literature do appear to infect the majority of host plants they were inoculated onto, even accounting for unreported failed inoculation attempts. For example, isolates of tobacco ringspot virus (TRSV, subgroup A) infected 290/361 species (80%), tomato ringspot virus infected 175/254 (69%, subgroup C), tomato black ring virus infected 174/208 (84%, subgroup B), Arabis mosaic virus (ArMV) infected 152/170 species (89%, subgroup A), and cherry leaf roll virus (CLRV) infected 117/137 species (85%, subgroup C). This last case is a virus of woody plants, which have caused issues in nepovirus inoculations before, such as blackcurrant reversion virus (Lemmetty et al., 1997, Latvala et al., 1997).
These are the only viruses with such a broad suite tested. Artichoke Italian latent virus (subgroup B) was tested against 91 (80% success) species and artichoke yellow ringspot virus (subgroup C) against 83 (92% success), for example.
There does appear to exist a possible difference between host range of the viruses as well, taking into account the assumptions made (see Discussion). Taking each plant species as a random variant around the imaginary mean of 'all plant species', with each species within a subgroup representing one observation of the subgroup, the chance of each subgroup entering any given plant is highest for A, lowest for B, with C between the two. Though subgroup B has fewer observations (at 490), subgroup C has the most (828) and subgroup A is close behind (815). This statistical test was provided by Roy Macarthur and the R script he provided is shown in Appendix A. While the limitations of the data mean that such an inference can't be considered concrete, it is an interesting hypothesis that the three subgroups differ in host range. It may form the basis for a follow-up study to investigate if this statistical inference actually points to an underlying difference in range and thus ecology for the different subgroups.

5.3. Discussion

5.3.1. Host Range of Nepoviruses

A table of host range of the members of the genus *Nepovirus* is presented here in Appendix B. This table includes data from host range studies, databases and textbooks, including where possible experimental or natural hosts (precedence given to experimental), hosts with demonstrated nematode transmission, and whether a challenged host was confirmed non-infected or merely asymptomatic.
Nepoviruses are considered easily mechanically transmitted to a wide host range (Hily et al., 2021). The assertion that nepoviruses have a 'broad host range' is difficult to quantify as the term is likely to be relative, but for over half of the known nepoviruses, inoculation attempts were reported for less than 40 plant species. *Orthotospovirus tomatomaculae,* in comparison, is stated to be species with a 'broad host range' and members of the species infect over 900 host species (Pappu et al., 2009).
From these data, it can be seen that *Nicotiana* and *Chenopodium* species are particularly amenable to inoculation, frequently occurring in the literature as purification or maintenance

hosts. C. *quinoa* is remarkable for being susceptible to infection from all 34 nepoviruses that it was inoculated with. The single failed reported inoculation comes from is a paper where symptoms did not develop after sap inoculation, including after back-inoculation to cucumber (Thomas and Procter, 1972). It is notable that this paper rubbed infected sap onto carborundum-dusted leaves with no buffer, as preparation of inoculum in phosphate buffers is known to increase success of inoculation (Hull, 2001). That is another facet to consider with the table: inoculation methods differ between studies. For example carborundum is no longer used at the laboratory where this thesis was conducted due to safety concerns; celite is instead used. The type of buffer and the recipe may differ as well.

Bias The uneven distribution of host-virus combinations could be reflective of a bias towards easily-manipulated viruses. Should a nepovirus prove amenable to mechanical inoculation, it may be tested against more viruses, or more specifically, such attempts may be reported in the literature. A brief attempt with a standard inoculation suite that does not infect any of the plants tested may go unreported, or at least mean that follow-up inoculations are not attempted as they might be considered a poor use of resources. The literature can be biased towards symptomatic hosts in even well studied viruses and hosts. For example, asymptomatic infection of rice with cucumber mosaic virus was not reported in the literature for some time after it was discovered (Roossinck and García-Arenal, 2015).
Not all studies report unsuccessful inoculations. A propagation host may be reported for the reader's use but no mention made of any attempts to inoculate onto a different host that were not successful, nor the number of attempts required before a successful infection was achieved. In some cases, asymptomatic hosts may be reported as non-hosts without serological or nucleotide-based tests, or back inoculation. In the case of back-inoculation, low titres can result in inconsistent success/virus recovery when back-testing (Jones et al., 1983), while still potentially being positive, preventing another layer where a positive host-virus association might exist. In terms of back-inoculation onto the original host, a study of cassava brown streak disease, virus isolated onto *N. benthamiana* was then unable to be back-inoculated onto cassava (Bock, 1994). Based on this, Were et al. in 2004 speculated on a mixed infection in their own samples when they did successfully induce symptoms on cassava during back-inoculation (Were et al., 2004). A study on citrus psorosis virus failed to back-inoculate the virus onto pepper likely due to challenging the host at the wrong time; the pepper was being used as an intermediary host to demonstrate Koch's postulates (Levy and Gumpf, 1991).
This failure to distinguish between non-hosts (immune), hosts not challenged, and hosts which may have a hypersensitive response, for example, has been noted before as a limitation on host range studies for diagnosis (Matthews, 1993).

Changes in nomenclature Confusion can also occur from changes in species names. Andean potato calico virus was first identified as a strain of *Nepovirus nicotianae* (tobacco ringspot virus, TRSV) in *Solanum tuberosum* (Fribourg, 1977), but is now assigned to the species *Nepovirus solani* (potato black ringspot virus, PBRSV), a closely related but separate species. For several

decades, beet ringspot virus was considered a strain of *Nepovirus nigranuli (*tomato black ring virus, TBRV*)*; it is now a member of its own species, *Nepovirus betae* (beet ringspot virus, BRSV) (Harrison, 1958, Lefkowitz et al., 2018). The Scottish TBRV-S isolate is now considered to belong instead to the species *Nepovirus betae*, while the two are distinguished by being vectored by different nematodes: *Longidorus attentuatus* (TBRV) and *Longidorus elongatus* (BRSV) respectively (Fowkes et al., 2021a). A study of viruses in fruit trees used commercial antisera to TBRV to detect this virus via ELISA, but found no *L. attenuata* when investigating the soil - only *L. elongatus* and *Xiphinema* species. The following hosts are listed as hosts of TBRV but may be hosts of beet ringspot virus instead or as well as: *Juglans regia, Malus pumila, Ribes nigrum, Prunus persicum, Rubus fruticosus* (Šubíková et al., 2002). 'Myrobalan' was also excluded as it is unclear to which species that Šubíková is referring to in the TBRV paper.

Disparity between natural and experimental host ranges In some cases, viruses may naturally infect a species but resist mechanical inoculation. For example, tomato is a known crop host of tobacco ringspot virus (EPPO, 2017) but was not successfully recovered from tomato inoculated from infected gladiolus and 'green strain' virus in experimental conditions (Bridgmon and Walker, 1952). ArMV infects white clover naturally, but could not be experimentally transmitted (Adams and Antoniw, 2006, Thomas and Procter, 1972).
Pelargonium hortorum is a natural host of TRSV and can be mechanically inoculated (Sastry et al., 2019), but was not experimentally infected by mechanical inoculation in another study even after back-inoculation as confirmation (Wingard, 1928). *Brassica rapa, Medicago sativa, Petrosilenum crispum, Solanum tuberosum* and *Pastinaca sativa* are natural hosts of TBRV (Converse, 1987), but not an experimental ones (Adams and Antoniw, 2006). *Ageratum conyzoides, Senna occidentalis* (Abraham et al., 2021) *Trifolium repens* (Puffinberger and Corbett, 1985) and *Capsicum annuum* (Gonzalez-Franco et al., 2014) are natural hosts of ToRSV but was not found to be mechanically transmissible.
Daucus carota was found to be infected with TRSV in the wild (Rush and Gooding, 1970, Powell et al., 1984) but attempts at mechanical inoculation of *D. carota* var. *sativa* have not succeeded (Price, 1940). As not all papers distinguish between wild and cultivated carrots, *Daucus carota* varieties been listed together, but it is worth noting that wild and cultivated carrots have been shown to differ in their susceptibility to mechanical transmission. Indicating variety and origin of host species is important for this reason (Van der Want et al., 1975).
Cases such as these may be due to variety of host, strain of the virus, or low transmission rates with small sample sizes. Not all papers report the number of plants per species challenged with the virus. In addition, transmitting viruses from *Rosaceae* hosts to other species is known to be interfered with by tannins and other secondary metabolites, where the inhibition is best overcome by use of specialist buffers (Hull, 2001).
Conversely, mechanical transmission needn't suggest that a virus might transmit naturally, or in crops grown in the field. In some cases, viruses can be mechanically transmitted where nematode transmission has not succeeded (Douthit and McGuire, 1978), or vice versa (indicated on the table). Similarly, a claim that a plant isn't susceptible based on lack of symptoms may

miss cases of successful transmission that, for example, may be asymptomatic, especially as some nepovirus symptoms can be masked by high temperatures (Komuro and Iwaki, 1968). In some cases, nematode transmission was indicated by association between a nematode, the virus in question, and the host in natural conditions, such as *Xiphinema americanum* and robacco ringspot virus in spearmint, watermelon, grape, apple and blueberry (Griffin et al., 1963, Fulton, 1962, Stone et al., 1962, Lana et al., 1983, Gilmer et al., 1970). In some cases, nematode transmission under experimental conditions is believed not to represent field conditions and may not indicate 'real-life' epidemiology (Harrison, 1977). Thus nematode transmission as indicated on this table may not always mean that this is likely to occur in a real outbreak.

Vectors In terms of the possible vector of the novel nepoviruses from oca, the name *Nepovirus* comes from two biological properties: the polyhedral shape of the virions and the frequent nematode transmission (Fuchs et al., 2017). Nematode transmission in nepoviruses is usually specific (Harrison, 1977), determined by a portion of the capsid protein (Hily et al., 2021). It is difficult to say for certain whether all members are nematode transmitted but there is evidence that some may not be (Hull, 2001). Whether this means that the virus is not transmitted by nematodes at all or simply not by the species of vector studied or under those conditions is unclear. However, CLRV has been tested for transmission by several *Longidorus* and *Xiphinema* species and is considered not to be vectored by nematodes (Jones et al., 1981). Blueberry leaf mottle virus is also deemed unlikely to be vectored by nematodes, based on an experiment where Xiphinema americanum individuals from Michigan and Arkansas were hand-picked to inoculate *C. quinoa* and *N. clevelandii* plants and did not produce a positive result (Brunt et al. 1996). One marker of nematode transmission used to infer a nematode vector is "patchy" distribution within a field (Gallitelli et al., 2012).

Ongoing viruliferous nematode infestations may cause re-infection over several growing cycles, minimised but not prevented by fallowing (Klingler and Zwicky, 1981). In the case of grapevine fanleaf virus (GFLV), weed hosts are unlikely to be significant to the epidemiology of this virus (Digiaro et al., 2017). GFLV along with other grapevine viruses are spread long-distance mainly by trade in infected plants (Digiaro et al., 2017).

Nepoviruses often display a high degree of vector specificity. Even the closely related GFLV and ArMV vectored solely by one species each, *X. index* and *X. diversicaudatum* respectively, which is determined by only 11 amino acids in the coat protein (Schellenberger et al., 2010). Cherry rosette virus was excluded as it is only tentative, has only one known host and was never assigned a subgroup, but was found to be vectored by *Longidorus arthensis* (Brown et al., 1998). Blackcurrant reversion virus is vectored by mites (and as far as is yet known, only by them), with artichoke yellow ringspot virus (AYRSV) a suspect for mite transmission due to its epidemiology and lack of other confirmed vectors (Susi, 2004, Maliogka et al., 2006, Karapetsi et al., 2021). Both are subgroup C. TRSV is transmitted weakly by grasshoppers (Dunleavy, 1957). Thrips can transfer n*I*-infected pollen from one plant to another (Dijkstra et al., 1998) with nymphs of *Thrips tabaci* able to transmit TRSV in soyabean at low rates (Messieha, 1969). TRSV is also reported to be transmitted by the spider mites in the genus *Tetranychus* (except *T.*

urticae (Granillo and Smith, 1974)), tobacco flea beetle and two species of aphids, all under experimental conditions (Adams and Antoniw, 2006).

A small number of nepoviruses have been demonstrated to transmit through pollen (Bhat and Rao, 2020), though this is of dubious ecological significance in the wild (Harrison, 1977) except in non-nematode transmitted species like CLRV and blueberry leaf mottle virus (Digiaro et al., 2017). Cycas necrotic stunt virus is is presumed not to be vectored either by nematodes or mites due to its widespread nature, but may rely on pollen transmission like CLRV is likely to (Shaffer et al., 2022). *Nepovirus* as a genus also contains the second greatest number of species demonstrated to transmit by seed, after *Potyvirus*, a trait that likely exists as an adaptation to the slow (30-50cm /yr depending on species is typical (Harrison, 1977)) movement of their nematode vectors (Sastry, 2013b).

PBRSV is a subgroup A nepovirus of potatoes thought to be common in Peru, but its vector is not yet known (Fribourg, 1977, Jones, 1978). Importantly, this is a nepovirus reported in oca in the literature (Lizárraga et al., 2000). Arracacha virus A is a subgroup A nepovirus of arracacha common in Peru, again with no clear vector (Jones and Kenten, 1978). Potato virus B (subgroup B) was first isolated from Peruvian potatoes and also has no confirmed vector (De Souza et al., 2017). Potato virus U (subgroup C (Adams et al., 2018a)) was first isolated in Peru and was reported to be transmitted by *Longidorus* nematodes collected from Peru (Jones et al., 1983). Tomato black ring virus (subgroup B) can be found in potatoes in Europe, with different strains vectored by *L. attenuatus* and *L. elongatus* respectively (Jones, 1978).

Reports of nematodes in oca include the discovery of *Thecavermiculatus andinus* (now genus *Altodera*), a heteroderid nematode (Golden et al., 1983), *Nacobbus abberans* and *Meloidogyne spp*., for which control is likely to be via resistant cultivars (Jatala, 1988). These nematodes are shared with ulluco, and some with yacón and arracacha (Jatala, 1988). A study of nematodes in the peridermis and parenchyma of different Andean root and tuber crops including ulluco and oca found a number of nematode species from different families, but none were longidorids (Lax et al., 2008). Longidorids are ectoparasitic (Neilson et al., 1999) and might have been removed with the peeling of the roots, but the study found *Criconema*, also ectoparasitic root feeders (Cordero et al., 2012). Any free-living soil nematodes may not have been detected, as the study aimed to investigate nematodes being traded with tubers, not the soil diversity.

The closest known relative to oca nepovirus 1 is cherry leafroll virus which, as noted above, has no known nematode vector. Whether ONV1 is transmitted horizontally by longidorid nematodes or some other means would require further study.

Blueberry leaf mottle virus and cherry leaf roll virus are both pollen-transmitted viruses vertically and horizontally; in perennial crops, this may be especially devastating as an entire population of plants may become infected over time this way (Boylan-Pett et al., 1991, Card et al., 2007). Though not a close relative of these nepoviruses, the oca strain of arracacha virus B (now understood to be a related genus, *Cheravirus*) is pollen transmitted in potato plants both vertically and horizontally, infecting 2% of potato plants grown from true seed of potatoes pollinated by infected pollen (Jones, 1982). Seeing that both closely related viruses and a virus

that shares a host, oca (which does produce sexual flowers (National Research Council, 1989)) are transmitted via pollen, either or both of these novel viruses could be pollen transmitted. Oca is a perennial plant (National Research Council, 1989). Even when dug up at harvest, volunteers from previous years are to be found in fields in the Andes, some of which are selected and grown to incorporate into landraces (Bonnave et al., 2016). ONV1 was tested for in many plants and seems to have a relatively low incidence. In fact, it was only ever found in a small number of plants from one source, a Polish seller on eBay. A number of studies on nepovirus transmission did only find the virus in the roots in some cases after nematode feeding, though in some instances this may have been due to contamination (Jones et al., 1981).

5.3.2. Experimental Host Range of Oca Viruses

The initial host range for inoculation with ONV1 was chosen based on a literature search, specifically hosts of its closest known relative, cherry leaf roll virus, as well as nepoviruses more broadly from a literature review. Some of the hosts chosen have a broad 'virus range' as well, as evidenced by this table (Christie and Crawford, 1978). ONV1 is only around 54% similar in sequence to CLRV however, so the relationship is quite distant.
Nonetheless, the broad host range reported for well-characterised nepoviruses was not evident in the case of ONV1. Over 100 individual plants were inoculated over several months, accounting for differences in season, stochastic sampling, and temperature. Even in the case of *C. quinoa*, infected in every nepovirus it was challenged with except for one under sub-optimal experimental conditions, was not infected. The only positive observation was one individual *C. sativa* plant, which, based on subsequent tests, was likely to be a non-indicative event. It is not clear why this individual plant was infected and no others, but it is not likely that the infection event represents a likely occurrence for this species-host combination and should not be factored into the epidemiology of the virus.
Per virus, per host, this can mean hundreds of plants tested for each novel virus. This may be costly in person hours and glasshouse space, as well as any follow-up tests to confirm asymptomatic infection. Given that this thesis' finding of many novel viruses within a relatively small number of samples chapter 3, a thorough host range test on a plant with even a very wide 'virus range' such as *Chenopodium quinoa* will be a hefty commitment if performed for every virus so found.

Need for an experimental host for oca viruses In the 2017 framework for assessing the plant health risk of viruses discovered via HTS, it was suggested that the choice of host might differ based on the desired outcome: in particular, that cultivars of the same species might be used to obtain a working population for biological study (Massart et al., 2017). Oca presents a difficulty with this approach: oca is not particularly well understood as a crop under experimental conditions. It has not been sequenced in its entirety. It does not possess many of the qualities which I summarised in the introduction as desirable for an experimental host. While I was able to determine that a number of plants were likely to be virus free, they are not readily

mechanically inoculated, in part because of their small leaves, in part because under heat or water stress the leaves rapidly fold, making them yet more difficult to inoculate. It is not known whether oca leaves contain phenols and other secondary metabolites which might interfere with the process of inoculating some viruses, or detecting them. What's more, though it is reasonably certain that some of the plants are virus free, the same cannot be said of other pathogens. Oca is known to be infected by black rot and charcoal rot, for example (Icochea et al., 1995, Sato et al., 1999), with stem rot reportedly common in North American plantings (https://www.cultivariable.com/instructions/andean-roots-tubers/how-to-grow-oca/). HTS data suggested the presence of *Rhodococcus fascians* in samples pooled together, extracted and sequenced as the dataset FR-BLK (French origin), which might explain the frequent symptoms of fasciation (Crespi et al., 1992). Determining symptoms on a plant whose response to stress and whose community of other infectious agents is not well understood is not straightforward with so many confounding factors involved.

Mutation It was thus desirable to isolate the virus into a different host for long term maintenance, though this isn't without its drawbacks. Maintaining virus in such experimental hosts can rapidly accumulate point mutations and would appear to be the main source of defective RNAs, which are probably less common in nature (Budzyńska et al., 2022). Muller's ratchet, aka 'mutational meltdown', whereby deleterious mutations accumulate in such fast-mutating organisms to the point of collapse of a population, is more common in populations with heavy bottlenecks, which can occur from single legion isolations (Clarke et al., 1993). One notable example of this genetic drift in nature was the frequent co-infection of *N. glauca* lines with tobacco mosaic virus (TMV) and tobacco mild green mosaic virus (TMGMV) over nearly 100 years in New South Wales, whereby one virus was lost entirely because TMGMV likely suppressed the accumulation of TMV and pushed its population below the threshold where mutational meltdown became likely (Fraile et al., 1997). Outside of the realm of plant viruses, serial passaging of a virus within sandfly cell lines led to a sixfold decrease in competence of vesicular stomatitis virus to infect mammalian cells, but after just one mammalian passage, the wild type functionality was restored (Novella et al., 1995, Roossinck, 1997). This is likely due to existing variation that was selected for during the passaging, and such bottlenecks may confound phylogenetic studies and analysis of ecological niches (Kuzmin et al., 2009). It is likely that the homogenous populations of host plants used to study viral evolution may not reflect the real-world complexity of a natural infection (Roossinck, 1997). Indeed, natural populations of viruses are extremely stable over time, compared with the examples above (García-Arenal et al. 2001).

Serial passaging local lesions from *C. quinoa* five times caused the infection rate on *N. tabacum* to drop by 30%, from near 100% transmission (Lafforgue et al., 2024). Another study found that repeated passaging through the local lesion host *C. quinoa* reduced virulence of Hibiscus chlorotic ringspot virus on *Hibiscus cannabinus,* probably due to a reduced ability to suppress gene silencing (Liang et al., 2002, Meng et al., 2006). In both these cases the critical factor is the lack of systemic infection leading to low MOIs. Another study in *C. quinoa* used intentional

mutagenesis of zucchini yellow mosaic virus and found that some mutants could infect *C. quinoa* but not the systemic host squash (Goh et al., 2023). While these mutants were created by laboratory mutagenesis, a naturally occurring mutant could plausibly be isolated this way during serial passaging, for example. Lastly, in the case of carrot yellow leaf virus, an isolate inoculated from carrot onto *N. benthamiana* was not able to be inoculated back into carrot, though it did infect chervil and coriander (Van Dijk and Bos, 1985, Van Dijk and Bos, 1989). This may have ecological relevance; viruses that escape from managed crops into the natural environment (spillover) may be attenuated when those viruses find their way back to the crop host (spillback) (Roossinck and García-Arenal, 2015).

Woody hosts also cause difficulties in back-inoculation; black raspberry necrosis virus was isolated only with difficulty onto herbaceous hosts and then could not be inoculated back into *Rubus sp.* (Jones and Roberts, 1977). Relatedly, ONV1 isn't the first nepovirus to be recalcitrant to mechanical inoculation onto herbaceous indicators. Inoculation of BRV from blackcurrant to herbaceous indicators took 'more than 20 separate attempts' (likely greater than the 100+ plants from this thesis, confirmed by Satu Latvala, pers comm) of which only one *C. quinoa* specimen was infected, from which other herbaceous indicators were later more easily inoculated (with this process being described with some bathos as 'with difficulty') (Lemmetty et al., 1997). This difficulty was suspected to be due to the unusually erratic distribution of the virus within the (woody) host as well as properties of the host cells, such as inhibitors (Latvala et al., 1997). Oca, however, is not a woody plant. The ease of transmission of ArMV from oca to the same indicator plants suggests that the difficulty in transmitting ONV1 is not due to any property of the host itself.

Comparison with highly virulent viruses The potexvirus was not inoculated onto as many hosts, nor was a table of hosts drawn up. Rather, the standard range was attempted with a small number of plants. It is still somewhat surprising that the virus did not transmit as potexviruses are transmitted by mechanical contact in nature (Hull, 2001). Contact transmitted viruses may transmit simply by leaves touching one another, such as the potyvirus, zucchini yellow mosaic virus (Coutts et al., 2013). This disparity is similar to the case with the genus *Tobamovirus*, which contains highly damaging members that are easily transmissible mechanically but is increasingly understood to contain latent members thanks to HTS studies (Ilyas et al., 2022). Standard host range tests with small (less than 10) numbers of plants like those attempted at the beginning of this thesis may be reliable for plants with known host ranges and high rates of mechanical transmission (such as ArMV, which transmitted reliably from oca to the same hosts). This is perhaps not viable for persistent lifestyle viruses or those otherwise not amenable to transmission, such as poleroviruses, which can complicate biological studies (Hoffmann et al., 2001). The focus on viruses of crop plants, and those that are easily experimentally manipulated, may have obscured the variety of the viral world (Roossinck, 2010).

The inability to transmit any of the viruses from the frozen material may speak more to trying to resurrect four-year-old samples than the infectivity of the viruses themselves. In particular, the tubers were frozen whole and, in leaf, whole frozen leaves lost infectivity (measured as local

lesion count) more than minced leaves or sap in the case of tobacco necrosis virus (Bawden and Pirie, 1950). ONV1, however, was from fresh material (picked and ground immediately) each time, as was the novel potexvirus.

Possible false positives and negatives in literature The second consideration is that symptomatic plants after inoculation needn't mean the desired virus was transmitted. The potexvirus-inoculated plants inoculated showed symptoms in some cases (of ArMV), but were negative for the potexvirus. These same plants were negative when tested via ELISA for ArMV; it was only qRT-PCR that indicated they were infected, and said qRT-PCR testing occurred because of the HTS data which indicated the presence of ArMV. A study on the host range of Cycas necrotic stunt virus reported positive, asymptomatic infections in *Beta vulgaris* and *Gomphrena globosa* based on chlorotic local lesions developing on *C. amaranticolor* after back-inoculation (Kusunoki et al., 1986). This back-inoculation/back-indexing method has been used to confirm asymptomatic infection in a number of nepovirus studies, e.g. (Ouertani et al., 1992, Hensen and Stace-smith, 1971). Though some studies used electron microscopy on the original plants used for inoculation or the infected plants (e.g. as above, (Ouertani et al., 1992, Hensen and Stace-smith, 1971)), viruses with similar morphology (e.g. members of the same genus) would be readily confused for one another, such is the case in nepoviruses in grapevine (Russo, 1985). For example, one study used electron microscopy on two different ultracentrifugation bands from infected material, where particles from one band had a more evident hexagonal structure than the other (Hensen and Stace-smith, 1971). In grapevine infections, nepoviruses were less apparent in electron microscopy than the same virus infecting indicator hosts such as *C. quinoa* (Russo, 1985)

Mixed and latent infections Mixed infections are quite common in nature, as revealed in recent years by metagenomics studies (Hasiów-Jaroszewska et al., 2021). There is a possibility that overlooking mixed infections may have resulted in an artificially inflated artificial host range for certain viruses when using back-inoculation as the method of diagnosis.
'Latent' or 'asymptomatic' infection with viruses can cause measurable yield losses, when taken at the population level (see (Budzyńska et al., 2020, Torrico et al., 2018, Nancarrow et al., 2022, Valentova et al., 2022)). Studies that take into account an entire population when determining causation can help overcome this symptomatic/asymptomatic confusion, such as association of carrot yellow leaf virus with necrotic symptoms (Adams et al., 2014), thus achieving two outcomes with one test. A study of citrus tristeza virus infections suggested differences in spectral reflection between control and infected populations, likely indicating stress in what was an otherwise asymptomatic population (Afonso et al., 2017).
This brings into question what it means for a virus to be latent or asymptomatic. The term appears to mean lack of apparent, visual symptoms in an individual (as the above papers use 'asymptomatic' in terms of plants with measurable yield changes). Even in visible symptoms, however, 'stunting' can only reasonably be inferred from comparison with other members of the species which are not stunted. A person seeing a stunted plant as their first specimen of a plant

species might assume that is the correct size of the species. The threshold between yield reduction and stunting is not always numerically delineated.
In the case of oca, such a comparison is difficult as there are no genetically homogenous, pest-free lines available as with potato to make such fine-tuned comparisons.
In plant viruses, founder effects can be quite strong, with very few virions initiating infection or participating in inter-cell movement (Lafforgue et al., 2024). The other novel nepovirus, ONV2, was detected via HTS as one small contig in total RNA extracted from plants which had been growing in the glasshouse between 2019 and 2021, suggesting that the infection was ongoing at detectable levels in plants for several years. Oca plants may appear as volunteers between croppings, some of which are integrated back into the cultivated crops if they have promising characteristics (Bonnave et al., 2016). These could act as a viral reservoir even when oca is not the active part of the cropping rotation.

Implications of low success This difficulty of transmission could result from a lack of purifying selection within a long-living, vegetatively propagated plant. This could lead to a population that is not under selection to transmit horizontally at high levels in order to thrive. For example, endornaviruses are dsRNA viruses found in plants and other taxa which are present in very low copy numbers and, apart from *Vicia fava endornavirus,* do not affect host phenotype (Fukuhara, 2019). They probably do not encapsidate and are spread solely via efficient vertical transmission with a persistent lifestyle, and their conserved RdRp is closest to ssRNA closteroviruses rather than ssDNA viruses (Fukuhara et al., 2006, Roossinck et al., 2011). They have a uniquely modular evolutionary history compared to other plant viruses and their lifestyle can be said to work with, not against, the plant defence system (Roossinck et al., 2011, Fukuhara, 2019).
One theory for the origin of this family is a derivation from a 'defective' ssRNA virus (Gibbs et al., 2000). The 'reduction' theory of viral origin proposes that gradual loss of extraneous functions may be what shaped the evolution of *all* biological entities with the viral lifestyle, analogous to function loss in other intracellular parasites (Claverie, 2020). In blunt terms, 'use it or lose it' may apply to plant viruses which, as a whole, possess very compact and efficient genomes.
HTS may be uncovering populations of virus which are in the intermediary state, members of taxa that are typically regarded as pathogenic (such as nepoviruses) but were previously unexplored, or both.
In the wild, many viruses are asymptomatic, as contrasted with those traditionally discovered by investigating virus aetiology (Roossinck, 2015). Generally, vertical transmission has been shown to favour reduced virulence; where host fecundity is the main means of a virus being transmitted, damaging the host is not favoured (Pagán et al., 2014). In some cases viruses may be mutualistic with their hosts (Roossinck, 2015). A well-known example is infection of panic grass with a fungus which is itself infected with a virus allowing thermal tolerance, and only when both virus and fungus are present (Márquez et al., 2007). In a study of cryptic viruses in blackgrass (*Alopecurus myosuroides*), Alopecurus myosuroides partitivirus 1 caused plants to develop many

tillers, with a shorter height, while Alopecurus myosuroides varicosavirus 1 had a less deleterious effect on the host when the host was under stress (Azcona, 2019).
On one hand, simply *being a virus* does not implicate the entity in causing disease. Nor does presence of a virus when a disease is present indicate the the virus is the cause of the disease, hence the need for demonstrating causation. On the other hand, lack of association with symptoms, yield loss, or other loss of fitness doesn't mean that a virus is irrelevant from an epidemiological standpoint, either now or under increasing climate destabilisation (Jones, 2009). Viruses asymptomatic in one host may be virulent in another, or may emerge from one host onto another, thereby leading to damaging epidemics (Roossinck and García-Arenal, 2015).

High-priority pests In cases of emerging viruses with severe economic consequences, characterisation may be accelerated. In the case of tomato brown rugose fruit virus, the virus' host range, epidemiology, host-virus molecular interactions and breeding efforts for resistance were already reported between 2015 and 2023 (Salem et al., 2023). Indeed, preventing an outbreak requires fewer resources in the long term than adaptive measures once an outbreak has occurred (Rodoni, 2009). For novel viruses which were found via HTS, however, biological information often remains limited (Hou et al., 2020, Rivarez et al., 2021). This limited understanding may be considered as an 'acceptable' unknown for many viruses which may not be high risk. However, a baseline of information is still essential to assert that the risk is indeed low and quantify any uncertainty (EFSA Panel on Plant Health, 2010).
A framework proposed by Fontdevila Pareta et al. in 2023 (to be used alongside and in support of conventional risk assessments) suggests leveraging public datasets and studies of field incidence to infer causation and epidemiology before undertaking more costly studies, such as experimental host range (Fontdevila Pareta et al., 2023). Sequencing of isolates, including from wild and reservoir hosts, may help to determine likely invasion routes and intervention points, e.g. seed in the case of bean yellow mosaic virus (Wylie et al., 2008, Rodoni, 2009). The findings within this chapter support the reasoning of Fontedevila Pareta et al., especially when considered together with the other chapters in this thesis.

5.3.3. Conclusion

In this chapter, a novel nepovirus was inoculated onto a suite of indicator hosts, including over 100 cucumber plants, based on a binomial distribution to determine the number of plants needed to siggest that the lone successful inoculation was atypical. Arabis mosaic virus was successfully transmitted from oca plants to herbaceous indicators under the same conditions.
Generally the data presented here are intended to serve as a commentary on the literature itself. For example, the statistical test showing a clear differentiation between host range of the three subgroups of nepovirus. Should this reflect a real difference in range, this would be a valuable insight into the ecology of nepoviruses. It is certainly worth further experimentation to confirm. However, even if this difference is due to the literature and the caveats above, this is itself interesting. It would indicate a severe bias in the literature, such as under-reporting of failed

inoculation attempts, lack of PCR confirmation and bias towards certain, highly damaging species such as ArMV and TRSV. A test on all 758 plant species for 41 nepoviruses would take considerable person-time to achieve (at 31,119 combinations, to say nothing of the case in this chapter where only 1/130 plants were positive), an endeavour which would require implausible funding for trained personnel to carry out. However, a representative suite could be determined and tested between the subgroups.

It has been proposed that, since virus' phenotype is determined by their genome, the limiting factor on our understanding of virus biology is not necessarily one of time or effort but of our ability to infer phenotype from genotype (Simmonds et al., 2017). Regardless of whether virus genomes are truly deterministic in this way, which is not yet known, there are limits to modelling future plant-virus interactions based on limited or incomplete historical data. This is especially likely to be the case for viruses with overlapping reading frames, where each nucleotide actually encodes for two separate proteins - in the case of PLRV, one reading frame appears to be subjected to stronger selection pressure than the other (Guyader and Ducray, 2002). Modelling both reading frames at once would likely be a considerable logistical challenge. In addition, the presence within a host of multiple genotypes of a virus with different properties is likely to result in complementation, where variants which are less fit in some regards can still affect virulence because they share the cells with more fit variants that can compensate for their reduced fitness (García-Arenal et al. 2001).

The increasing quantities of viral sequence data provided by metagenomics have opened up the possibility of training machine learning algorithms on these datasets with applications for predicting subcellular localisation and RNA silencing suppressors from amino acid sequences, as well as folding predictions of protein products (Ghosh et al., 2022). However, these, while undoubtedly useful, focus on individual molecular interactions, and are a long way from predicting the complexities of virus-host interactions in their entirety.

At present machine learning is usually limited to 'traditional' models based on predicting a particular feature using labelled datasets which, in biology, is ideally compared to an empirical dataset in testing to determine its correctness before deployment (Greener et al., 2022). As summarised by this review in a flow chart form: "Sufficient data?" -> "No" -> "Get more data". At present the 'get more data' stage is still necessary if predictive modelling is to be utilised. These data are challenging to obtain at present. In the Discussion of this thesis (chapter 7 on Page 147) the possible implications of finding so many hard-to-transmit viruses is discussed in a risk assessment context.

Chapter 6. Investigation of Viruses in Transcriptomes from the NCBI Short Read Archive

6.1. Introduction

As public databases expand with transcriptomics data of cellular organisms, virus discovery has been extended to include sequences detected as contaminants in host transcriptome data, including plants, known as data-driven virus discovery (Lauber and Seitz, 2022). Data mining through tBLASTN and assembly of the Transcriptome Shotgun Assembly archive (TSA) and SRA specifically searching for homology to known viruses have revealed a number of new members of the genus *Amalgavirus*, (Sidharthan et al., 2022b) the family *Secoviridae* (Sidharthan et al., 2022a, Verhoeven et al., 2023), the genera *Polerovirus*, *Deltapartitivirus* and *Enamovirus* (Kavi Sidharthan et al., 2022),(Kavi Sidharthan et al., 2023) the genus *Solendovirus*, several rhabdoviruses (Sidharthan and Baranwal, 2021, Bejerman et al., 2022, Bejerman et al., 2021), members of the order *Tymovirale*s (Bejerman and Debat, 2022) family *Barnaviridae* (Nibert et al., 2018) and astroviruses (Lauber et al., 2019). Mining of transcriptome data from an initiative to sequence 1000 plant transcriptomes (1KP) revealed many other RNA viruses, including in plants not typically well characterised in virology (Mifsud et al., 2022, Debat et al., 2023).

In terms of studies of viruses infecting particular hosts, examples include a study of the transcriptome of soybean at different developmental stages that was later used to link these developmental stages to viral presence (Choi et al., 2023) a study of the virome of Cabernet sauvignon grapes from a transcriptome study (Jo et al., 2018) and DNA viruses in the tree *Weltwitschia mirabilis* (Debat and Bejerman, 2022). Publicly available pepper transcriptomes were mined to specifically identify infections of the sampled plants with bell pepper endornavirus (Jo et al., 2016). A study re-analysing datasets used for virome analysis discovered a plethora of other plant pathogens from fungi to insects, some of which were later confirmed by qRT-PCR on the original sample, meaning that virological studies can be re-used for other disciplines as well (Haegeman et al., 2023). Such usage of publicly available datasets for survey work was proposed as a means of quickly ascertaining alternative hosts and geographic range of viruses in a 2023 framework for investigating viruses first detected via HTS (Fontdevila Pareta et al., 2023).

A recent innovation is the Serratus 'palmID' database, which scanned petabases of nucleotide information to identify viral RdRp sequences. These identified a 'palmprint' representing the hand-like conserved catalytic core of the RdRp, typically around 103 amino acids long (Edgar et al., 2022). This database (serratus.io) has since been used in a number of papers, for example

expanding known *Waikavirus* diversity (Sidharthan et al., 2023), finding of a novel luteovirus and a novel potexvirus from an HTS study in existing datasets of *Prunus* and papaya respectively (Khalili et al., 2023, Cabrera Mederos et al., 2022), identification of two flaviviruses (Debat and Bejerman, 2023) and detection of a novel ilarvirus in a variety of datasets (Rivarez et al., 2023a). Serratus was used in the detection of tomato fruit blotch virus as part of an aforementioned paper on rhabdoviruses (Blouin et al., 2023, Cabrera Mederos et al., 2022) and investigation of ophioviruses in the 1KP study (Bejerman et al., 2022, Debat et al., 2023). One study used the new database along with SRA datasets and obtained the genome of an isolate of Telosma mosaic virus by assembling reads from a public transcriptome of *Pogostemon cablin* (Gou et al., 2023). Lastly, the database was used in part for identifying grapevine fanleaf virus proteins for study (Roy et al., 2024).

This data-driven virus discovery shortcuts the need for gathering biological samples by leveraging the far greater number of studies sequencing host genomes or transcriptomes, that have nonetheless sequenced "stowaways", but suffers from similar drawbacks to other HTS methods, such as checks on assembly quality and difficulty detecting highly divergent sequences (Lauber and Seitz, 2022). Such analysis also has regulatory concerns, as viruses found via SRA data could be linked to trade sanctions if it is found in a country that exports related plant products (Macdiarmid et al., 2013, Fontdevila Pareta et al., 2023).

This chapter further investigates the virome of what are novel commodities in the UK, as part of a preliminary risk assessment. Data were analysed from publicly available transcriptome data derived from sequencing of ARTC samples. Reads from the SRA datasets were downloaded using the SRA toolkit before bioinformatic analysis using a modified version of Angua. The resulting viral contigs were further manually investigated using Geneious Prime (see section 2.4 on Page 40).

6.2. Results

Only datasets derived from short read sequencing were considered (i.e. not datasets containing reads obtained from PacBio sequencing, Nanopore or related long-read technologies). Though no plant virus specific comparison has been made, it is possible that choosing only short read sequencing technologies may overlook some taxa; for example, a comparison of Oxford Nanopore and Illumina MiSeq for investigation of 16S rRNA in a mock community in soil found the reported abundance was influenced by the sequencing platform used (Lemay et al. 2022). Studies which utilised only DNA sequencing with no cDNA step were not used as they do not reverse transcribe RNA into DNA, thus the machine cannot detect inputted RNA sequence. They would thus overlook RNA viruses within the samples entirely, leaving out a potentially large part of the sample virome. 16S studies aimed at bacterial barcoding were likewise excluded for being DNA-only.

Table 6.1 A table of the SRA runs used for data mining, including the BioSample from which they originate, the geographic origin of the sample, and any associated papers.

SRA Accession	BioSample	Host	Origin	Variety	Sequenced	Paper
SRR7873465	SAMN10081146	Oca	Peru	Unknown	Fera, UK	Jones et al., 2019a
SRR7873467	SAMN10081141	Mashua	Bolivia	Unknown	Fera, UK	Adams et al. 2018a
SRR18215731	SAMN25047287	Yacón	Shenzhen, China	Unknown	Chinese Academy of Agricultural Sciences	Fan et al. 2022
SRR18215732	SAMN25047287	Yacón	Shenzhen, China	Unknown	Chinese Academy of Agricultural Sciences	Fan et al. 2022
SRR18215740	SAMN25047287	Yacón	Shenzhen, China	Unknown	Chinese Academy of Agricultural Sciences	Fan et al. 2022
SRR10244774	SAMN12993632	Yacón	Poland	Unknown	Fera, UK	Silvestre et al. 2020
SRR797211	SAMN01984102	Yacón	Peru	Unknown	University of British Colombia, Canada	Hodgkins et al. 2014
SRR2912199	SAMN04251315	Maca	Yunnan, China	Unknown	Zhengzhou Tobacco Research Institute	Chen et al. 2018
SRR3440095	SAMN04901587	Maca	Yunnan, China	Black	Zhengzhou Tobacco Research Institute	Chen et al. 2018
SRR3329550	SAMN04901486	Maca	Yunnan, China	Black	Zhengzhou Tobacco Research Institute	Chen et al. 2018
SRR7003713	SAMN08932292	Maca	Yunnan, China	Unknown	Chinese Academy of Forestry	Shang et al. 2018
SRR7003714	SAMN08932294	Maca	Yunnan, China	Unknown	Chinese Academy of Forestry	Shang et al. 2018
SRR2960160	SAMN04293597	Maca	Yunnan, China	Unknown	Kunming Institute of Zoology	Zhang et al. 2016
SRR2960161	SAMN04293597	Maca	Yunnan, China	Unknown	Kunming Institute of Zoology	Zhang et al. 2016

The SRA accessions, associated BioSamples, geographic origin and associated papers of the samples are summarised in Table 6.1 on Page 139. In total fourteen SRA datasets from eleven different BioSamples, one mashua, one oca, six maca, and three yacón plants were investigated. These originated variously from China, Poland, Bolivia and Peru, with some SRA datasets representing samples derived from preserved historical isolates.
Some ARTC samples originate from the same institution as this thesis does, Fera Science Ltd. In the interest of only using public datasets, however, they were downloaded and analysed as if they originated from an outside institution. These were the sole oca sample, the sole mashua sample, and one yacón sample.
The findings of viral sequences in these samples are summarised in Table 6.2. No sequences of novel virus candidates were assembled from the reads of any datasets analysed.

Table 6.2 A table of the SRA accessions and the viral sequences detected when analysing them. Percentage identity of the recovered sequence (nucleotide) to known viruses is given as well as the approximate proportion of the genome recovered by all matching contigs, per segment if the virus is segmented.

SRA Accession	**Host**	**Genus**	**Virus**	**& nt identity** (region)	**% Genome Recovery** (segment)
SRR7873465	Oca	Cheravirus	Arracacha virus B	99 (all)	Complete (RNA2), 62 (RNA1)
SRR7873467	Mashua	Potyvirus	Mashua virus Y	94-100 (all)	100 (monopartite)
			Papaya ringspot virus	94 (polyprotein)	3 (monopartite)
			Wild potato mosaic virus	93-99 (polyprotein + 3' UTR)	9 (monopartite, fragmented)
			Potato virus V	99 (CP)	3 (monopartite)
		Carmovirus	Carnation mottle virus	97 (RdRp + CP)	20 (monopartite, fragmented)
		Orthotospovirus	Chrysanthemum stem necrosis virus	1 (RdRp)	2 (RNA L)
		Carlavirus	Potato virus S	99 (RdRp)	4 (monopartite)
SRR18215731	Yacón	Tobamovirus	Cucumber green mottle virus	100 (RdRp)	3 (monopartite)
		Allexivirus	Garlic virus A	99 (NBP)	3 (monopartite)
SRR18215732	Yacón	Badnavirus	Yacon necrotic mottle virus	95 (ORF1-4)	65 (fragmented)
SRR18215740	Yacón	Badnavirus	Yacon necrotic mottle virus	96 (ORF1-4)	53 (fragmented)
SRR10244774	Yacón	Badnavirus	Yacon necrotic mottle virus	96 (ORF3)	4 (monopartite)
		Ilarvirus	Potato yellowing virus	99 (all)	96.7 (RNA1), 100 (RNA2), 97.4 (RNA3)
SRR797211	Yacón	Ilarvirus	Potato yellowing virus	95 (RNA1 ORF)	100 (RNA1), 0 (RNA2), 0 (RNA3)
SRR2912199	Maca	Potyvirus	Bean yellow mosaic virus	98-100 (polyprotein)	5 (monopartite, fragmented)
SRR3440095	Maca	Potyvirus	Bean yellow mosaic virus	100 (polyprotein)	2 (monopartite)
SRR3329550	Maca	Potyvirus	Bean yellow mosaic virus	98 (polyprotein)	7 (monopartite, fragmented)
SRR7003713	Maca	Potyvirus	Turnip mosaic virus	96 (polyprotein)	81 (monopartite, fragmented)
SRR7003714	Maca	Potyvirus	Turnip mosaic virus	94-98 (polyprotein)	86 (monopartite, fragmented)
			Sugarcane mosaic virus	99 (all)	95 (monopartite, fragmented)
SRR2960160	Maca	Polerovirus	Brassica yellows virus	88-95 (all)	100 (monopartite, fragmented)
		Potyvirus	Turnip mosaic virus	85-97 (all)	93-100 (monopartite, fragmented)
SRR2960161	Maca	Polerovirus	Brassica yellows virus	88-98 (all)	100 (monopartite, fragmented)
		Potyvirus	Turnip mosaic virus	88-96 (all)	100 (monopartite, fragmented)

6.2.1. Oca and Mashua

Only one mashua plant on the SRA archive was sequenced via extraction of RNA that was reverse transcribed into cDNA, though a small number of studies originated from 16S amplicons. This RNASeq dataset of mashua, SRR7873467, originates from Fera Science Ltd. and was stored as a Bolivian historical isolate of MVY (Adams et al., 2018c). Along with the mashua virus Y which was the intended target of the initial sequencing run, a number of small contigs showed homology to other viruses.

The mashua dataset provided sequences matching 98.8-100% nt identity to papaya ringspot virus W and P, which are the watermelon and papaya strains (Quemada et al., 1990). The two contigs matched different regions of the potyviral polyprotein coding sequence, and may be contaminants at only 3% coverage, or else misidentified portions of the MVY genome. One 366bp contig matched the RdRp of potato virus S (PVS) at 97% identity. PVS has previously been reported in arracacha and potato, but not mashua (Brunt, 1990, Santillan et al., 2018b, De Souza et al., 2018). The close ecological relationship between the hosts, however, could make this a genuine detection rather than contamination, expanding the number of viruses shared between two or more ARTCs, but it is not possible to be certain of this finding at such low coverage.

The following sequences were also detected: a 240bp contig matched potato virus V with 99.8% identity, a 211bp contig that matched the orthotospovirus Chrysanthemum stem necrosis virus (CSNV) with 100% identity, and three contigs between 269 and 411 bp in length matching to wild potato mosaic virus with 90-100% identity. Most of these findings are potyviruses. Thus, even if they are not contamination, there is a possibility that they are recombinant or divergent sequence from within the MVY population in the host. Tomato spotted wilt virus, an orthotospovirus, was reported in European ulluco samples in 2022 (De Jonghe et al., 2022). The carmovirus carnation mottle virus was also detected, with 20% genome recovery.

As with mashua, the only oca-derived HTS dataset on the SRA was from extraction and sequencing of a historical isolate, this time of a sample of oca leaves infected with the oca strain of arracacha virus B (AVB-O). Sequence homologous AVB-O was recovered from this sample, albeit not the entire genome as was found in the original paper (Jones et al., 2019a). No other sequences showed homology to viruses.

6.2.2. Maca

Two of the datasets of RNA sequenced from extraction of maca plants, SRR2960160 and SRR2960161, originated from the same BioSample, SAMN04293597A. A number of contigs with similarity to turnip mosaic virus (TuMV) were chimeric in origin from both of these datsets. These are marked with a star in Table 6.2 as any divergence in sequence from the reference sequence may be due to this issue and must be interpreted with caution. However, of the reads which are certainly not chimeric, all are clearly closely related to TuMV, with the entire genome recovered across the various fragments. Where coverage is given as a range, this range is excluding chimeric reads (host/virus and virus/virus) that contain part TuMV segments, and

including them in the TuMV portion respectively. For example, one sequence assembled that showed homology to the TuMV coat protein contained sequence matching the glycoprotein gene in its 5' terminal, which is not the correct order for TuMV, but the rest was correct. Removing the entire read reduced the coverage, though the sequence was represented within the contig. For this reason, both numbers are given in the table. It is not clear why the potyvirus resulted in so many chimeric contigs only when assembling reads derived from this one BioSample. However, for the purposes of diagnosis, it is clear that the maca plant sampled as SAMN04293597A, as well as the BioSamples SAMN08932294 and SAMN08932292 were at the time of harvesting likely infected with TuMV.

6.2.3. Yacón

Some sequences derived from assembly of reads from the yacón dataset SRR10244774, from a sample collected in Poland, showed close homology to PYV (>97% identity). As with the sequences of PYV derived from HTS sequencing of samples from the Irish supplier (subsection 4.2.3 on Page 100), the PYV sequences from assembly of the SRR10244774 dataset suffered from some misassembly errors. Accounting for these errors, the sequence from the yacón was very similar to Yacon-York isolates (MN527502-4, >98% identity), as would be appropriate given that the NCBI isolate 'York' was sequenced from RNA extracted from the same sample of yacón (BioSample SAMN12993632). Detections of sequence with high homology to accessions of Dahlia mosaic virus were excluded from the table as they come from a known endogenous strain (see section 6.3) but were quite abundant in all yacón datasets.

6.3. Discussion

Yacón and maca samples are represented most often in the data. The datasets from HTS sequencing of their extracts originate mostly from Chinese institutions; this is presumably due to the recent commercial importance of yacón and maca in China (see section 1.5 on Page 10). All oca and mashua datasets originated from stored historical isolates of viruses infecting ARTC hosts. These isolates were subjected to RNA extraction and sequenced in the UK from Andean samples. Most ARTCs were not represented at all in transcriptomic data.

6.3.1. Oca and Mashua

The assembly of the HTS dataset originating from sequencing of the historical isolate of AVB-O (SRR7873465) did not recover the entire genome as the originating paper did. For the purposes of surveying for presence, however, this finding was sufficient to confirm that AVB-O was present in this oca sample even without the associated paper (Jones et al., 2019a) based on genome coverage. In terms of using a public dataset for confirming findings in other hosts and areas, this would be sufficient.

The contigs obtained from assembly of the single mashua-derived dataset were mostly very short (as little as 2% of the genome in the case of Chrysanthemum stem necrosis virus), making it

difficult to ascertain if there is a genuine infection or if the reads assembled originated from contamination. The exception is the carmovirus carnation mottle virus (CarMV). Assembly of the reads from sequencing of the mashua transcriptome led to recovery of 20% of the genome of CarMV. CarMV was first reported in the UK (Brunt, 1996). In the past CarMV infection was almost ubiquitous in carnation plants and infected 33/113 test species, but was at that time presumed to only naturally infect *Dianthus* species (*Caryophyllaceae*) (Hollings and Stone, 1964). CarMV is still extremely widespread and found everywhere carnations and sweet williams are grown, including Chile and Argentina, though virus-testing and elimination programs have been reducing its incidence (CABI, 2024). The CABI Compendium also lists ornamental plants in other families as hosts, such as *Araceae, Begoniaceae, Gentianaceae, Orchidaceae, Thymeleaceae* and the crop *Lactuca sativa* (*Asteraceae*, lettuce) (Zheng et al., 2011, Chen et al., 2011). This natural infection in mashua would be the first report in *Tropaeolum sp.*, meaning that a targeted PCR test to diagnose CarMV in mashua would be worthwhile to confirm the finding and include mashua in the host range of CarMV.

6.3.2. Maca

The findings from assembly of reads from the SRA suggest that maca is a host brassica yellows virus and turnip yellows virus (which are likely to be one species (Filardo et al., 2021)) and potyviruses such as bean yellow mosaic virus and TuMV. Turnip yellows virus is an important pathogen of oilseed rape, first identified in the United Kingdom and vectored by aphids (Filardo et al., 2021, Duffus and Russell, 1970). Infection with turnip yellows virus is regarded as the major cause of the yield gap in UK oilseed rape (Stevens et al., 2008), and the virus is widespread in UK pea crops (Fowkes et al., 2021b). Even apparently asymptomatic turnip yellows virus infections were found to cause yield losses of as much as 40%/28% in peas/lentils (*Lens culinaris*) respectively when experimentally infected via *Myzus persicae* in a field trial (Nancarrow et al., 2022).

Oilseed rape, a host of turnip yellows virus, is a major oil crop in China, which could explain the presence of turnip yellows virus on maca from this country (Bonjean et al., 2016). Though maca is traditionally grown at high elevations, availability of seeds in Australia at a small scale (Hermann and Bernet, 2009) suggests that the crop can be grown at lower altitudes; the highest mountain in Australia is 2228masl (Verrall et al., 2023). Should seeds of maca become available in Europe the same way other ARTCs have been, they may act as an inoculum source and contribute to viral pressure on commercial oilseed rape and legume plantings. This is not likely to be a large concern however, as turnip yellows virus already has a large number of reservoir species in the UK, which contribute to renewed infections each year (Slavíková et al., 2022).

TuMV was also detected in the maca plants. TuMV was reported as a natural infection in Yunnan maca in 2015 as the CHN12 isolate (Yin et al., 2015). TuMV infects abroad range of hosts, so it is perhaps unsurprising that it infects maca, another member of the brassica family (Tomimura et al., 2003). was mechanically inoculated onto maca before, in 2003, so the status of maca as a potential host was known before the report of a natural infection (Marthe et al., 2003). TuMV

causes economically important damage in brassicas worldwide (Shattuck, 1992) and is thought to have originated in Europe and radiated to Asia from there (Tomimura et al., 2003). Likely, then, the maca plants sequenced in China were infected with TuMV outside of South America. SRA accession SRR7003714, a dataset submitted from Yunnan, China, when assembled, generated contigs matching with greater than 99% identity to an accession of sugarcane mosaic virus from maize in China (accession MT519600). The sequences of sugarcane mosaic virus recovered from these maca plants represent 95% of the genome, and thus are not likely to originate from contamination. Sugarcane mosaic virus is a threat to both Asian and European maize production, and cereals more broadly (Wu et al., 2012).
It is unknown if any of the viruses detected in these datasets cause deleterious symptoms in maca. The detections of these viruses do indicate that maca is being exposed to these economically important viruses and appears to be successfully infecting maca plants in the climate in which it is currently grown in Yunnan. At present it is not known if there is interest in a European market for maca seeds, but none were found for sale during the present study.

6.3.3. Yacón

As with reads derived from sequencing RNA from the yacón samples from the Irish commercial supplier (subsection 4.2.3), sequence with homology to geminivirids was frequently found in the yacón SRA datasets studied here. As mentioned, such viral sequences from DNA viruses are a known confounding factor when investigating plants via HTS, as HTS alone cannot distinguish integrated from active infections (Brait et al., 2024). The 'Dahlia mosaic virus strain D10' contigs all match to virus which is known only to be endogenous: Dahlia variabilis endogenous plant pararetroviral sequence (DvEPRS) (Pahalawatta et al., 2008, Maricle and Gillock, 2020). The very small sequences with homology to yacón necrotic mottle virus could come from an active infection, even a symptomatic one, but the HTS data cannot indicate this. That is one of the drawbacks of data-mining SRA reads in place of a large scale survey where samples are collected and processed with express intent to detect viruses.
PYV is a virus that is known to infect yacón, including in Europe, and may in fact have evolved with yacón before later expanding its host range to include potato (Silvestre et al., 2020). PYV is an A1 quarantine pest (EPPO, 2023a). These findings add to the evidence that PYV is widespread in yacón plants, including those being grown in Asia.

6.3.4. Conclusion

Here, data-mining of SRA datasets from transcriptome studies of ARTC samples has been used in support of risk assessment for a commodity pathway. The lack of association, let alone casual, with symptoms, is one drawback of data-driven virus investigations which have very minimal biological data associated with them. Any information must be extracted from associated papers, which may have different priorities from a virological study and thus miss details which a virologist might need to make an assessment, such as symptom expression or possible vector presence.

The short length of many of the contigs with homology to viruses assembled from these reads also represents another limitation of using public SRA data for surveying. It is not always clear from SRA metadata which datasets may have been sequenced together with other datasets, i.e. which samples were multiplexed on the same lane as one another, if any. Likewise it is not always indicated which positive and/or negative controls were used, making a comparison with controls or other hosts on the same sequencing lane difficult. This information can be used to determine the likelihood of contamination, for example. Based on this fact, none of the very small, lone contigs found in these data should be taken as confirmation of a virus within the plants studied as the possibility of contamination cannot be excluded. The tentative detection of virus sequence in these datsets may, however, indicate that a targeted test for these viruses in ARTC hosts from that location, or HTS aiming to sequence viruses within said hosts, may be advisable. For example, knowing that a short contig with species-level homology to PVS sequence was detected in mashua, which originates from the same cropping system as arracacha and potato, a targeted assay to detect PVS in mashua, such as PCR, could be undertaken.
The advantage of data-mining the SRA for virus detection comes from the method's low cost and relatively high speed. Provided there is computing power and expertise available, a survey of viruses within SRA data may be conducted at potentially little cost. Availability of cloud services such as Galaxy may allow smaller labs without dedicated servers to conduct such a survey at lower cost than maintaining a dedicated computing infrastructure (Massart et al., 2014). In cases of niche crops such as ARTCs, however, such a survey is limited to a relatively small number of samples as the transcriptomes of these crops are not frequently sequenced, and thus few public database entries of HTS reads are available in order to investigate for viral sequences.

Chapter 7. Discussion

7.1. Introduction

The aim of this thesis was to provide a case study of the risk of internet trade in plant propagules to UK plant health. A case study of Andean root and tuber crops was chosen, for being plants for planting which are widely available through e-commerce in the UK, often from mainland European sellers. Viruses were focused on, using HTS to detect known and unknown viruses in the samples collected. The virome of ARTCs was investigated through several means.

The first method, covered in chapter 1 used a literature search to establish a list of known viruses known to infect ARTCs in South America, as well as reports from more recent studies which investigated viruses infecting ARTCs grown as niche crops in Europe.

The second method utilised HTS to detect viruses infecting oca tubers available through internet purchases in order to determine the risk of internet trade of ARTCs to European plant health (chapter 3). Total RNA was extracted from these tubers which was then sequenced using HTS in order to identify both known and unknown viruses present within the samples. Further oca samples and one mashua sample were purchased and surveyed using qRT-PCR for the presence of four novel viruses detected in the HTS datasets. While ONV1 was limited in incidence to only the batch of plants it was first isolated from, one virus, OAV1, was found to be widespread and present in samples from several countries. When other *Oxalis spp.* samples were tested for presence of same four novel viruses, some naturalised *Oxalis spp.* plants growing in urban environments tested positive for OCV1, suggesting that oca grown in Europe may be infected with European viruses. No quarantine viruses were detected in oca during the present work.

In addition, a comparison was made between the plant health status of oca plants purchased from poorly regulated internet trade with ARTCs sourced from a commercial nursery in Ireland with good phytosanitary practice (chapter 4). The oca plants from this supplier were mostly virus-free, except for RNA extracted from a pool of samples from five oca plants, which tested positive for beet ringspot virus (BRSV) when sequenced with HTS. This infection would require confirming via another method, such as qRT-PCR, as the positive result was not certain. The four yacón plants from this supplier were infected at a 50% incidence with the A1-listed virus PYV. In addition, further novel viruses were detected in RNA extracted from the single mashua sample obtained from the commercial supplier in Ireland. Fragmentary sequence from another virus, which may belong to a novel species, was detected in RNA extracted from a pool of five ulluco plants.

As part of biological characterisation, sap from ONV1-, OVX- and OAV1-positive plants was mechanically inoculated onto a series of indicator hosts (chapter 5). In particular, sap infected

with ONV1 was inoculated onto many hosts, with *Cucumis sativa* plants inoculated over 130 times. Of these plants, one tested positive for ONV1 by qRT-PCR. A literature search of nepovirus hosts was conducted and presented as a table. The findings of this search were summarised and it is suggested that subgroups A, B and C may differ from one another in their host range, but this requires further investigation. In addition, the data indicate a heavy bias towards inoculating nepoviruses onto a small number of host species, particularly from the families *Solanaceae* and *Amaranthaceae*,and for testing only a small number of viruses on many host species.
Lastly, ARTC transcriptomes available as datasets on the Short Read Archive were downloaded and their reads assembled and investigated for sequences indicating viral infection (chapter 6). A small number of these SRA datasets originated from sequencing of historical isolates of known ARTC viruses, such as AVB. While no novel viruses were detected, maca sequenced and sampled in Yunnan, China was found to be infected with poleroviruses and potyviruses. It is hypothesised that the poleroviruses and potyviruses present in the maca plants are Asian in origin, not South American. Other viruses of commercial importance were detected but due to the quality of the sequences recovered, some of these viruses cannot be said for certain to be infecting the plants. This demonstrates some of the limitations of re-using sequence data from non-virological studies for virus detection.
This discussion chapter brings together these findings, and is presented using the framework of an IPPC risk assessment in order to discuss their relevance to plant health, particularly in Europe, as well as to discuss possible limitations and difficulties of the PRA format when dealing with viruses as possible pests.

7.2. International Standards for Phytosanitary Measures

The aim of this study was to utilise HTS in support of risk assessment of a possible pathway risk, that of the unregulated internet trade in ARTCs. A number of potential pathogens, new and known, were identified in the plants being traded, confirming the hypothesis that trade in ARTCs over the internet is introducing novel viruses to the UK. It is likely that this pathway will be circulating the viruses found around Europe, and possibly globally. Further work was then undertaken to further investigate the overall risk posed.
Plant health risks under IPPC regulations are determined from a set of international standards of pest management (ISPMs) either presented by a pathway or a pest. These are ISPM s No. 2, (IPPC Secretariat, 2019) No. 11 (IPPC Secretariat, 2013) and No. 21 (IPPC Secretariat, 2021). The findings of the present work will be discussed based on how they fit to the guidelines of categorisation given by these documents. ISPM 2 considers identification of a pest or pathway of concern more broadly, and is linked to pest risk assessments (PRAs) for quarantine pests and regulated non-quarantine pests respectively. The difference between quarantine pests (ISPM 11) and regulated non-quarantine pests (ISPM 21) is that the former are absent from the region entirely, while the latter are known within the PRA area but present an unacceptable risk when associated with plants for planting (the plants for planting must be the main transmission

pathway of the infestation) (IPPC Secretariat, 2021). Under EPPO guidance, a plant protection organisation of which the UK is a member, quarantine pests are split into A1 and A2 pest list categories, representing pests absent from the region under EPPO surveillance and pests that are present but not widespread in the region respectively. There may be smaller pest risk areas defined individually per pest, such as individual countries or high-risk areas. These are worded as recommendations and may not apply to every individual member state (EPPO, 2023e). These two lists are analogous to pests determined by the risk assessments outlined in ISPMs 11 and 21 respectively. The differences between ISPM 11 and ISPM 21 are such that ISPM 21 does not consider market access, environmental damage or probability of establishment as the pests in question are already present within the area of concern; pests in this case are regulated insofar as they affect the use of the plants intended for planting (IPPC Secretariat, 2021). ISPM 11 states that 'A list of pests likely to be associated with the pathway (e.g. carried by the commodity) may be generated by any combination of official sources, databases, scientific and other literature, or expert consultation.'

Invasion risk as a concept can be assessed based on a taxonomic unit, on a pathway itself, such as a commodity imported for a particular use, or based on qualities of the site being invaded, such as conservation or ecological service value (Mcgeoch et al., 2016). ISPM 2 identifies this as initiation and this may begin with either a pest or a pathway but not explicitly based on the importance of the at-risk area itself (IPPC Secretariat, 2013). The present work was developed as a pathway investigation of internet purchases of the commodities under the umbrella of ARTCs. As far as the pathway is concerned, the summary of the steps touched upon are taken from the sub-sections of ISPM 11 (considering these pests as potentially quarantine) and are:

- 1.1 Initiation points
  - 1.1.1 PRA initiated by the identification of a pathway
  - 1.1.2 PRA initiated by the identification of a pest
- 2.1 Pest categorization
  - 2.1.1.1 Identity of the pest (here the HTS data and subsequent characterisation and phylogeny).
  - 2.1.1.2 Presence or absence in the PRA area (here, the survey of viruses present on the commodity and of plants related to the commodity)
- 2.2 Assessment of the probability of introduction and spread
  - 2.2.1 Probability of entry of a pest
    - 2.2.1.2 "occurrence of the pest in a life stage that would be associated with commodities",
    - 2.2.1.3 "speed and conditions of transport and duration of the life cycle of the pest in relation to time in transport and storage " and "prevalence of pest likely to be associated with a consignment".

        - 2.2.1.4 “Probability of pest surviving existing pest management procedures” generally (chiefly evading inspection)
    - 2.2.2 Probability of establishment
        - 2.2.2.1 Availability of suitable hosts (related also to the survey, and to the experimental inoculations in a looser sense)
            - Especially: “whether there are other plant species, which could prove to be suitable hosts in the absence of the usual host species”.
        - 2.2.2.2 Suitability of environment
            - Regarding mostly the properties of the host itself and its amenability to differing daylight hours.
        - 2.2.2.4 Other characteristics of the pest affecting the probability of establishment
            - With specific reference to genetic adaptability (here discussed in the lens of viruses broadly and the mixed infections commonly present in these samples).
    - 2.2.3 Probability of spread after establishment
        - With particular focus to: intended use of the commodity, its potential for movement within borders and across them.

- 2.3 Assessment of potential economic consequences
    - 2.3.1 Pest effects
        - 2.3.1.1 Direct pest effects
        - 2.3.1.2 Indirect pest effects
    - 2.3.2.1 Time and place factors
    - 2.3.2.4 Non-commercial and environmental consequences

This prsent study focused primarily on these outlined steps of the pest PRA process for a PRA initiated by a pathway (i.e. a commodity not previously imported into a country, in this case ARTCs). Here the pest risk area is defined as the UK, with some consideration to Europe more broadly. Specifically, the countries where these tubers were sourced from are under consideration: Ireland, Hungary, Poland, France, the Netherlands and the UK.
As noted in the literature review (chapter 1) and broadly by virologists, when one investigates a commodity (or any) plant via HTS, it’s likely that a member of at least one novel virus species will be sequenced, which will need to be assessed as a potential plant health risk (Macdiarmid et al., 2013; Olmos et al. 2018; Fontdevila Pareta et al., 2023). Any pathway assessment, then, can be considered as a commitment to investigate at least a handful of viruses, when HTS is used as the first step of investigation. This association between pathway risk assessment and pest categorisation is recognised in ISPM 2 (IPPC Secretariat, 2019). ISPM 11 states that: “. . . a

number of individual PRAs may be necessary for the various pests potentially associated with the pathway. The opportunity to eliminate an organism or organisms from consideration before in-depth examination is undertaken is a valuable characteristic of the categorization process." (IPPC Secretariat, 2013). EFSA further provides guidance on splitting these ISPMs into two phases: phase one is pest characterisation, determining if it should be further investigated for risk, and phase two is pest risk assessment via expert solicitation and modelling (Jeger et al., 2018). In terms of the present study, the flow is as illustrated in Figure 7.1. Though the flowchart, begins with sample acquisition, further sample inputs can be added at any stage of the process. These may include more members of the host species, ideally other varieties, weeds from around infected crops or known to be associated with the crop species, or related species. For example, with *Oxalis tuberosa*, which is not native to the UK, other *Oxalis* species grown as ornamentals or wild in the UK were sampled for the survey work to test for presence of the detected novel viruses within the UK and Europe more widely. This process is a useful option when a crop species is niche or otherwise not widely present in an importing country already and is recommended in the updated framework for characterising HTS findings (Fontdevila Pareta et al., 2023).

Determining potential weed reservoirs can assist in untangling the epidemiology of a given virus and has factored into a number of HTS surveys before (Hasiów-Jaroszewska et al., 2021, Rivarez et al., 2023b). Baselining in this manner has been recommended as a form of basic research to understand novel virus findings (Macdiarmid et al., 2013). Authors of a catalogue of viruses of indigenous species on Aoteoroa-New Zealand noted that utilising sequence data from these species on the SRA may expand the number of viruses known to infect organisms endemic to Aoteoroa-New Zealand (Robson et al., 2022). Assembly of reads from SRA datasets has already been utilised to expand the known host range of tomato fruit blotch virus and Solanum nigrum ilarvirus 1, for example (Blouin et al., 2023; Rivarez et al. 2023a). Investigation of public archives for the presence of sequences derived from EVEs can also expand the host range of virus taxa (Aiewsakun and Katzourakis, 2015).

It should be noted that the focus on pests over pathways often leads to retroactive assessment of an organism only after an outbreak has occurred and may overlook means of managing many unwanted organisms, including unknown ones, by regulating just one pathway (Evans et al., 2010). A reliance on individual pest risk assessments only after a pest has been recognised is a weakness of the present phytosanitary system (Brasier, 2005). For these data then, the overall risk of the pathway itself should be considered as more than the sum of individual organisms. Chiefly, this work demonstrates that plants for planting traded through small purchases on the internet is not well regulated, and indeed presents challenges to regulation that traditional trade does not. Consider, then, that the novel viruses here are only a selection of what may be circulating via couriers from e-commerce purchases in the UK, and that more work is certainly required to close the gap.

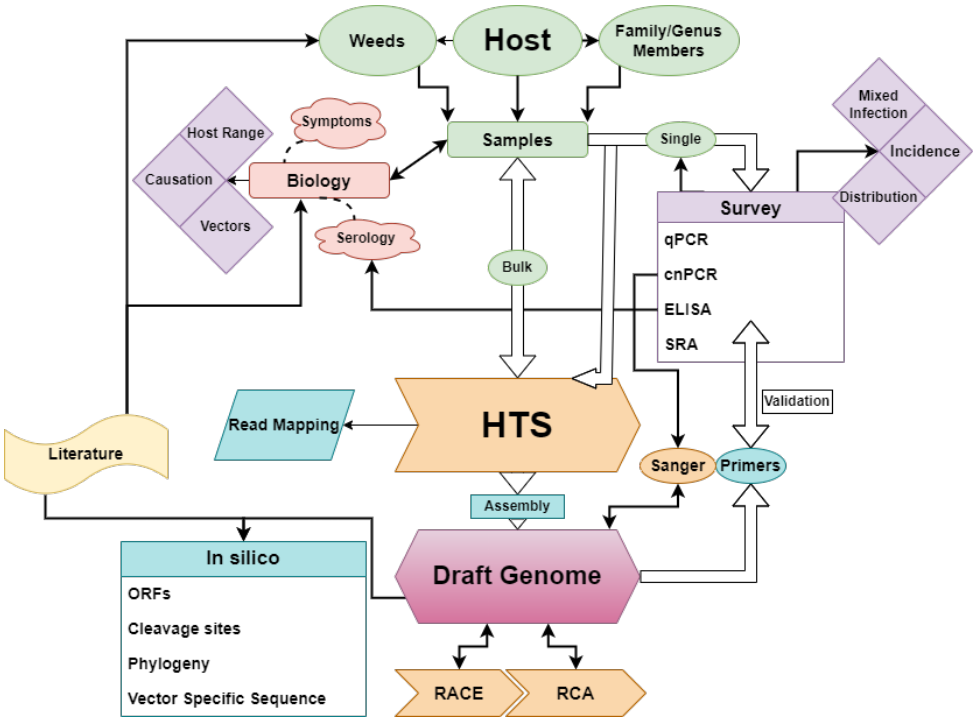


Figure 7.1 A flowchart of how the risk assessment process was undertaken in the present work. Large arrows indicate the 'major flows' of information integral to this thesis and smaller arrows indicate where additional clarifying information may be obtained. Sample acquisition begins when a novel pathway has been identified. Before any targeted tests begin, HTS is used to determine the virome of a pool of representative samples for economic reasons. Based on the dataset from sequencing these pooled samples, further work involves primer design, a survey of the virus within plant samples pooled to form the initial HTS sequencing run, a baseline of known and novel viruses in potential reservoir hosts, and literature searches to inform biological characterisation. Draft genomes may originate from pooled or individual samples, which may be further refined by Sanger sequencing, RACE, RCA and other molecular methods. The Short Read Archive can be mined for viruses if relevant transcriptomes are available.

7.3. 1.1 Initiation Points

To begin a pest risk assessment, the pest must be identified as such. A pest is defined by the IPPC as: “Any species, strain or biotype of plant, animal or pathogenic agent injurious to plants or plant products” (IPPC Secretariat, 2023b). The pathogen’s identity must be known or symptoms must be shown to be transmissible if the exact etiological agent is not known, though an organism’s membership of a known pathogenic genus might also be used to indicate that an agent is a pest (IPPC Secretariat, 2019). Once the pest status of an organism is known or suspected, the established presence or absence in the PRA area will determine whether a pest is quarantine or non-quarantine but regulated (Bragard et al., 2020, IPPC Secretariat, 2021).
Every putative novel virus discovered during the present study is proposed to belong to a genus with known pathogenic members. However, the use of HTS in plant virology has led to a recent and large expansion of some of these genera. Many of these new species have members which are not associated with symptoms or lack biological context beyond an isolation host. For example, many novel cytorhabdoviruses were described in just one study with no known associated symptoms (Bejerman et al., 2023). A similar study uncovered sequences from many ophioviruses which likewise are not known to be symptomatic (Debat et al., 2023). Generally, virus taxa down to even the species level are polythetic; sharing a common origin and lineage but not necessarily a “single common defining property” (Van Regenmortel et al., 2013). Nonetheless, the PRAs listed here would consider these novel viruses as members of pathogenic taxa and investigate them as such.
In addition to the novel viruses identified and described, relevant literature was consulted in (chapter 1) to determine viruses known to be associated with ARTCs generally, though not necessarily known to be associated with the host in Europe. Some of these viruses are on the A1 pathogen list as of February 2024 (EPPO, 2023a). A number of viruses discovered or detected in ARTCs (Table 1.3 on Page 16 and Table 1.2 on Page 15) are A1 pests. These are arracacha virus B and its oca strain on arracacha and oca (Jones, 1981, Kenten and Jones, 1979), potato black ringspot virus in oca and arracacha (Jones, 1981, Kreuze et al., 2020a), Andean potato latent virus and/or Andean potato mild mosaic virus on maca and ulluco (Kreuze et al., 2020a, Bragard et al., 2021), potato virus T in mashua, ulluco and oca (Lizárraga et al., 2000), and potato yellowing virus in yácon, including European samples (Silvestre et al., 2020). Blueberry leaf mottle virus was reported in oca (De Jonghe et al., 2022) but the virus detected was likely to have been oca nepovirus 2 based on examination of the data during this thesis (see paragraph 3.2.11). The one virus detected in ARTCs that is on the A2 pest list (EPPO, 2023b) is pepino mosaic virus detected by HTS in internet-traded European yacón (De Jonghe et al., 2022).
In terms of ’eliminating’ a pathogen or at least downgrading the assessment of its risk is made possible through a number of means and may not require extensive biological characterisation. This is especially important when so many viruses are being discovered in recent years with only limited biological data associated with them (Hou et al., 2020). A particularly useful method bypasses the need for intensive biological assessment by confirming that a virus is already likely

to be indigenous to Europe, thereby meaning that some interventions, such as exclusion of the agent from the area, are not necessary or even useful. For example, OCV1, described from oca, was detected in the UK infecting naturalised weeds and ornamental *Oxalis*. If this detection reflects possibility of wider spread of OCV1 within the relevant climactic zone of Europe it is, at worst, a regulated non-quarantine pest by this fact alone, and may not be of concern at all. In order to be considered for regulatory intervention as a non-quarantine pest, a pest must be demonstrated to cause an 'unacceptable' impact when associated with plants for planting (IPPC Secretariat, 2021). Surveys may help alleviate the burden of extensive characterisation in this regard by rendering the matter less urgent. The biological work performed for this thesis works as a cascading series of gates via which a pest might be eliminated for consideration as a quarantine pest. Thus, the cheapest and quickest methods for eliminating a pest ought to be tried first, saving resources for the viruses likely to represent a plant health risk in the UK and wider Europe.

7.4. 2.1 Pest Categorisation

7.4.1. 2.1.1 Identity of the Pest

Case studies Confusion related to virus identity has previously caused regulatory difficulty. Two virus species whose members have been detected infecting ARTCs can be used as examples here.

The first is BRSV. As noted in section 5.1, confusion relating to virus nomenclature, strains and species membership has already made the true experimental host range of some viruses somewhat vague. Identification of BRSV as a strain of *Nepovirus solani* and then its own species, along with recombination, have caused inconsistent regulation of this pathogen in regions where it is regulated (Fowkes et al., 2021a). BRSV, which is not quarantine in the EU or the UK, was first isolated in Scotland and is likely endemic in Europe (Harrison, 1957b).

Until 2019, the genome of PYV had not been sequenced in its entirety (Knierim et al., 2019). This lack of genomic information meant that, in some datasets, nucleotides sequenced from PYV isolates can be incorrectly labelled as Fragaria chiloensis latent virus (a close relative) if using BLASTN and BLASTX to identify contigs before this time. This is because the closest known homology on these pblic datasets was Fragaria chiloensis latent virus sequence, and not the as-yet-unknown PYV sequence. Fragaria chiloensis latent virus is not on either A1 or A2 pathogen lists, but PYV is an A1 pathogen (EPPO, 2023a, EPPO, 2023b). Accession KU375548.1 on NCBI, not attached to any paper, appears to originate from a Polish NPPO and (likely incorrectly) identifies Fragaria chiloensis latent virus in yacón plants. As PYV is an A1 quarantine pest in the EU and Fragaria chiloensis latent virus is not (EPPO, 2023a; EPPO, 2023b) the oversight is concerning, because it means that quarantine viruses were being incorrectly identified as a non-quarantine virus. The infected material was thus potentially not handled with proper quarantine procedures. Fortunately this confusion is no longer likely to occur as PYV is now well-characterised and homology-based characterisation is more likely to

detect the proper homology to isolates of PYV. Other emerging pathogens, however, in particular those whose genome has yet to be sequenced fully, may cause similar problems in future. These issues identifying viruses relate also to the difficulty of classifying viruses as biological entities. Virus taxonomy is the logical process of applying human categories to a class of organism which is very much inhuman (Van Regenmortel, 2007). Virus taxonomy differs from cellular life because there is no universally conserved gene or gene group on which to build a tree containing all viruses (Simmonds et al., 2017). Viruses as a class are polyphyletic and different phyla may have different origins, perhaps even within phyla, with attempts to reconstruct an evolutionary tree of even one Baltimore classification remaining difficult (Claverie, 2020).

Difficulties with viruses as a class Viruses, especially positive-sense ssRNA viruses and ssDNA viruses, recombine at especially high rates, which has had implications for molecular diagnosis if the sequence of an ORF used for diagnosis later turns out to be prone to recombination (Martin et al., 2011, Bentley and Evans, 2018). In particular, the following groups of viruses reported in oca are known to recombine with other viruses extensively both within and without of their species: nepoviruses, poleroviruses, capulaviruses (as with Geminiviruses more broadly), potyviruses, potexviruses, allexiviruses (Hily et al., 2021, Latourrette et al., 2021, Bernardo et al., 2016, Valli et al., 2007, Revers et al., 1996, Sherpa et al., 2007, Celli et al., 2019). Ophioviruses have been posited to reassort their segments in nature (Thekke-Veetil et al., 2015). That is to say, members of every genus to which the novel viruses described in this present work belong, except the two rhabdoviruses, have been reported as recombinant in some manner in the literature. Some virus species within these taxa may even originate from a recombination event, such as grapevine deformation virus, hypothesised to be formed as a recombinant of ArMV and GFLV (Elbeaino et al., 2012). Many virus species may trace their origin to such an event, termed symbiogenesis (Roossinck, 2005). Such recombination and reassortment challenges the ability to assign viruses and their sequences to species based on monophyly (Van Regenmortel et al., 2013).
A relatively small number of mutations can change biological qualities and lead to emerence of new strains and biotypes (Van Regenmortel et al., 2013). Even single nucleotide changes both natural and induced have been linked to differences in symptomatology in tobacco mosaic virus and in cucumber mosaic virus (Banerjee et al., 1995, Sleat and Palukaitis, 1992). However, while RNA viruses mutate quickly due to their error-prone replicase, it is likely that their proteins are no more variable than those of DNA viruses due to negative selection, and that viruses may, in general, conserve proteins to around the same degree as their hosts (García-Arenal et al. 2001). The use of EVEs to study ancestral viruses also suggests that viruses may evolve slower over extremely long periods of time than their short-term mutation rate might suggest, though whether this observation is an artefact of sampling bias or the result of alterations in biology and evolution rate over long periods of time remains to be understood (Aiewsakun and Katzourakis, 2015).

Multipartite viruses present their own challenge, as typified here with the novel ophiovirus(es). In short, the difficulty comes in determining which set of three sequences presented here 'belongs' to which virus (that is to say, which sets of three segments likely co-evolved together and whose sequences taken together represent an example genome of one species). The nepoviruses are also bipartite viruses but, fortunately, nepoviruses typically share a common sequence in their 3' UTR (De Souza et al., 2017) between the two genomic segments which helps with association of the two.
Multipartite viruses which encapsidate their genomic segments separately can reassort as well, further confusing the matter. An experiment combining the top and bottom fractions of viruses from two different ophioviruses of citrus ringspot virus and citrus psoriasis virus produced an infection, suggesting that these aryificaly reassorted viruses were competent within the plant cells (García et al., 1993). Ophioviruses are now known to naturally swap segments with one another and this process contributes to their evolution (Martín et al., 2006, Thekke-Veetil et al., 2015). Such reassortment presents a taxonomic impasse when describing what appear to be two or three different viruses within the same sample based on thresholds of sequence identity. In theory, the ophiovirus sequences detected in oca are likely to be from different species based on nucleotide and predicted amino acid identity, (see subsection 3.2.7). However, when it comes to epidemiology and biology, the reality of virus infection can be more complex than that. Even among members of different species within the same genus, the true existence of a multipartite lifestyle (i.e. separately encapsidated genomic molecules as opposed to different genomic molecules encapsidated together) can vary, and it requires experimental confirmation (Sicard et al., 2016).

Considerations of viral biology encountered in this thesis All novel species candidates presented in this thesis were identified using HTS, with little to no biological information for most beyond the host and some hints at geographical distribution. ICTV species demarcation criteria are assigned per taxon by the study group and generally use nucleotide and amino acid thresholds which maintain distinctions between groups of viruses with certain biological traits. Though genomic information alone is sufficient as of 2017, phenotype (and predicted phenotype) and ecology are taken into account if possible (Van Regenmortel et al., 2013, Simmonds et al., 2017). Whether some viruses with sequence homology represent several species within the same genus or one diverse species is at times unclear, especially where nucleotide divergence sits at the threshold, because it is not always possible to robustly demonstrate monophyly. This has regulatory import. A collapse of various bunyaviruses from 257 to 94 species without differentiating via strain or biology led to concerns that viruses with symptoms as broad as haemorrhagic disease in humans and nonspecific febrile illness might be confused for one another in regulatory and scientific circumstances (Blitvich et al., 2018).
For ONV1, which is presented here as master sequence derived from HTS of a pooled sample (paragraph 3.2.11), the sequence, while representative, is unlikely to represent any one genome and is likely a chimera of related genotypes from within the plants whose tissue was combined and extracted to form the pool (Simmonds et al., 2017). Fortunately, this pooled sample

contained sequence derived from plants that originated from the same purchase and likely the viruses within form one closely related lineage. Where possible only contigs that originate from one sample have been presented as draft genomes. Draft genomes that are sequence derived from RNA extracted from multiple plants were checked for congruence with sequence from other, separate samples.

Detecting a virus population and assigning it to a species, especially a species known from older literature without genomic information, also presents a challenge to determining risk of ARTCs. As noted in the literature review, serological relationships may well have masked different, closely related polerovirus species in ARTCs and potato before: these are Andean potato latent virus and potato leafroll virus, though in the former case there is a possible PCR confirmation in the literature (Lizárraga et al., 1996a, Lizárraga et al., 1999a, Fox et al., 2019, Bragard et al., 2021; Lizárraga et al., 1996b). Mashua virus Y was described as a new virus but is highly likely to be sequence belonging to a previously described Andean virus.

The novel potyvirus described from in mashua during the present work, TropPV3, could be truly novel or could be newly-detected sequence belonging to a previously described virus. This is relevant because re-describing the same virus erroneously increases the total number of viruses known to be associated with a commodity. More importantly, it changes the information available that can be used to infer the origin of a virus population. For example: is the Tropaeolum potyvirus 3 sequence representing a new species, or is the virus that was sequenced actually Tropaeolum potyvirus 1 or Tropaeolum potyvirus 2? As well as this, potyviruses in general have biological properties which can complicate study of their phylogenetic relationships with one another. Different species collectively referred to as the PVY subgroup of potyviruses have at times been regarded as strains of potato virus Y due to similarities in portions of their genomes (Bejerman et al., 2010, Inoue-Nagata et al., 2006, Sanches et al., 2014, Adams et al., 2005b). Based on coat protein sequence alone, TropPV3 might have been considered an isolate of Bidens mosaic virus, for example.

Associating sequence to known viruses is important both for biology and for tracking historical and present geographical distribution. It prevents doubling of work in characterising a virus and connects the two 'eras', combining benefits of HTS and older, more biologically complete experiments, commonly on viruses causally associated with symptoms (Jones et al., 2021).

Lastly, the mentioned complication of multipartite viruses was encountered in this thesis, with the novel ophiovirus(es). This difficulty in determining genomes of segmented viruses has been noted before as a limitation of assembly-based virus discovery (Simmonds et al., 2017). The sequences are represented, then, as four RdRp sequences, three CP sequences, and three MP sequences, one set of which is likely to represent a partial genome a one new virus species (see subsection 3.2.7). As stated, they may be different species or highly divergent variants of the same species, given their similar niche and predicted properties, but too little is known to be assured of this.

7.4.2. Causal Association in Plant Virology

ISPM 11 links identification of the pest, immediately, with a casual association between virus and symptom before a virus can be considered quarantine (IPPC Secretariat, 2013); as stated, this is how 'pest' is defined by the IPPC (IPPC Secretariat, 2023b). A virus being quarantine but not a pest would be a contradiction of terms, in theory. Causation thus acts as a fulcrum to the entire 'pest' risk assessment process.

Koch's postulates Fulfilment of Koch's postulates is a common method of demonstrating causation in plant pathology more broadly (Agrios, 2005). The triad of postulates as applied to cellular pathogens commonly seen throughout secondary literature is an adaptation, never expressed by Koch in quite so direct and concise a fashion; they are not universal but form guidelines for developing means associating organisms to disease casually (Ross and Woodward, 2016).

The triad of postulates seen in secondary literature has been adapted several more times to suit the disciplines of plant pathology and plant virology, including molecular variants (Bos, 1981, Falkow, 1988). Here I will discuss the findings of the viruses in this thesis as they apply to the observations posited by Fox in 2020 (Fox, 2020). What should be noted though, is that the applicability of Koch's postulates to virology has been in question since nearly the beginning, before the nature of viruses was especially well understood (Fox, 2020, Rivers, 1937). In particular Rivers draws attention to the interaction of plant viruses: "... at least one natural disease of plants is induced by the combined action of two viruses, each of which has been obtained free from the other and [...] each produces a characteristic malady different from that caused by the synergistic action of the two agents". A well-known and early example of this is the case of the potyvirus potato virus Y and the potexvirus potato virus X, where the two were separated via the means of inoculation: aphid or needle (Smith, 1931).

A plant virus (or other plant microbe) being 'a pathogen' is not a quality of the organism itself but is situational, indeed able to be managed and ameliorated in some cases by human interventions (Jeger et al., 2021). What it even means for one virus to be one virus, for one organism to be one organism, becomes more blurred the more we as humans understand the biological world (Dupré and Guttinger, 2016). In this case, associating any of the one-to-three novel ophioviruses in this thesis to symptoms would be impossible as it is not even certain which one is which. The novel allexivirus was 'associated' with symptoms but it was also 'associated' with a plant which was in a parcel for several days prior to being observed. That is to say, there is no basis for assuming a causal relationship between the observed mottling of the leaves and the viral infection.

This relates also to the inability to transmit any viruses here through mechanical means. In this study infectious clones were not produced, though they may have had superior results for obtaining an infection than mechanical inoculation did (see chapter 5). This would not have been possible due to lacking a portion of the nepovirus genome, but the complete sequence of other viruses were available for cloning, in theory.

Infectious clones Infectious clones can provide a 'pure' isolate where mixed infections might be difficult to separate out through means of differential hosts, especially where viruses involved might have a wide host range, allowing for biological studies (Fox, 2020, Massart et al., 2017). Infectious clones can allow viruses to transmit into a plant without the helper virus they would require in nature, such as umbraviruses (Taliansky and Robinson, 2003) (though there have been cases of an umbravirus being detected without its helper, such as pea enation mosaic viruses -1 and -2 in *P. sativum* in the UK and Germany, possibly due to mechanical inoculation in the field (Fowkes et al., 2021b, Gaafar et al., 2020)). One such study enabled a deeper understanding of the molecular biology of a population of opium poppy mosaic virus (Ilyas et al., 2021). The naming of the two viruses reported in the initial paper suggests that it is the umbravirus causing the symptoms and the polerovirus which is demoted to being associated when, in reality, either or both could be responsible for the symptoms observed on poppy and herbaceous indicators (Tang et al., 2016). Mechanical inoculation of the infectious clone of the umbravirus by itself was asymptomatic on *N. benthamiana*, for example (Ilyas et al., 2021). Since opium poppy mosaic associated virus (subsection 4.2.4 on Page 102) may be at least in part responsible for the disease symptoms on *P. somniferum*, it is one of the few viruses reported in this thesis that is associated (though not *causally*) with symptoms.

In terms of causation, utilising infectious clones to demonstrate Koch's postulates may be a case of finding methods to fit a framework, rather than finding a framework which suits the organisms. For example, in 1937 Rivers noted that maintaining virus in tissue cultures was already being used as a means of demonstrating the 'pure culture' third portion of Koch's postulates in viruses, and notes that Koch could not plausibly have had such a thing in mind (Rivers, 1937).

An infectious clone may represent just one haplotype amplified via PCR, such as by amplifying overlapping fragments in a plasmid or restriction enzymes, often with primers designed to a consensus/master genome sequence (e.g. (Kondo and Fujita, 2012, Pasin et al., 2018, Pasin et al., 2019)). In mastreviruses and other DNA viruses, clones may be generated by amplifying DNA from purified virions and digesting it into a construct, removing the need for genomic information to be known and thus for primers to be purchased (Boulton, 2008). In RNA viruses, clones may be cDNA (Pasin et al., 2019). In this way, infectious clones are removed by varying degrees from the true viral population which was present in the original living host.

Infectious clones typically require addition of a promoter and some artificial means of introducing the clone into the host (Bhat et al., 2020). A human finger rubbing extracted sap in phosphate buffer onto a wounded plant is hardly likely to happen in a field or commercial glasshouse either, but what it is using is an extant sub-sample of the viral populace. It begins the process at the (typically encapsidated) 'virion' stage of the life cycle as would happen in a natural infection, as opposed to beginning from synthetic sequence.

In terms of obtaining multiplying populations of virus in an easily-maintained plant species, either method can be useful. Infectious clones provide excellent model systems for studying molecular biology *in planta* and have even been used to 'resurrect' ancient virus (Pasin et al.,

2019). For demonstrating causality though, other experimental or observational means may be more appropriate than infectious clones.
Similarly, symptom development after back-passaging in a species of plant with a known deficiency in its RNA silencing system (*N. benthamiana*) (Todesco and De Felippes, 2016) may not be a strong case for it causing disease in crops without such genetic quirks, which tend to be cultivated to resist viral disease, not to be susceptible to it (Kobayashi et al., 2014).

The disease triangle Hill refers to the complexity in the following manner: "the whole chain may have to be unravelled or a few links may suffice. It will depend upon circumstances" (Hill, 1965). The monocausal model of disease states that one agent is both *necessary* and *sufficient* for disease development (Broadbent, 2009). In nature, it is not always a simple case that the presence of a disease's casual agent will inevitably lead to development of that disease. Factors such as host resistance and tolerance, cross-protection (Goh et al., 2023), temperature and/or darkness (Kassanis, 1957), drought (Mishra et al., 2022), mixed infections and other factors, biotic and abiotic (McLeish et al., 2019), all affect the development of disease once a virus is replicating inside a plant. This complexity is reflected in the modern disease triangle of plant pathology (pathogen, environment, host) (Agrios, 2005). In some cases this is extended with a fourth axis of human involvement, which may even render the pathogen as the 'proximate cause' and humankind as the 'ultimate cause' due to poor management of cropping systems (Zadoks, 2001). Koch's postulates are often interpreted in a manner which is monocausal, though from an interventionist perspective, they do allow isolation of a causal agent on which it is possible for us as humans to act upon and exclude to prevent disease (Ross and Woodward, 2016). Determining a casual relationship between organism(s) and disease, and associated factors making the disease more likely, will vary for each pathosystem and must be investigated according to the needs of the situation at hand. This has implications for risk management. Quarantine, where successful, is an intervention which excludes the 'necessary' portion of the pathosystem altogether (in those rare cases where such a requirement is absolute), but effective biosecurity has its own costs and externalities pre-and post-border (Burgman et al., 2013). In some cases it may be more appropriate to apply leverage in other areas of the pathosystem to improve the efficiency of resource allocation. These fall in line with other successful methods of demonstrating causality that work along gradients or continua of time, dose response and consistency, taking into account polymicrobial causes and other environmental factors (Hill, 1965, Fox, 2020).

7.4.3. 2.1.1.2 Presence or Absence in the PRA Area

From phylogenetic information and survey data it is possible to infer that some of the viruses infecting oca available for sale on the internet in the UK are European in origin, though whether they are widespread is more difficult to determine. For example, OCV1 is likely to be already widespread in Europe and infects other *Oxalis sp.* as well as *T. thailctroides*, but the other

viruses may be limited only to oca at present. ONV1 in particular probably has a more limited distribution in Europe.

Surveys A survey of endemic plants, especially those with a close taxonomic relationship to the commodity in question, can help to determine presence or absence of a virus in the area under consideration. For example, the detection of OCV1 in naturalised *Oxalis sp*. plants growing as weeds in the South of England, and/or in ornamentals for sale from UK-based sellers would suggest that the virus is already present in the country. Possibly, OCV1 first infected the oca plants sampled (or a direct ancestor) at some point in the recent past. It is, however, not impossible that OCV1 is South American in origin and has escaped onto European oxalids from introduced oca which were infected in their native Andean region. This distinction is not so easily made with some other viruses, which may have their origin in South America, Aoteoroa-New Zealand, Europe or even North/Central America, with no data which might lead to a hypothesis that favours one origin or the other.
In particular, oca belongs to a family with no other genera included but *Oxalis*. Other ARTCs, such as yacón, might prove more amenable to a virological investigation by surveying for viruses in taxonomically related plants. As an aster, yacón belongs to a family containing a great number of horticultural and agricultural crops grown in Europe, including many agricultural weeds (Weber and Gut, 2005). In the case of mashua, the plant is related to nasturtiums, such as *Tropaeolum majum,* the garden nasturtium, which as a native of Peru and was a popular ornamental in the Incan Empire (National Research Council, 1989). *T. majus* is now a naturalised or invasive plant in parts of Europe (Duenas-Lopez, 2022). *T. majus* hosts a broad range of viruses which may threaten its close relative, mashua (Grau et al., 2003). In the present work, a virus first reported in poppy and nasturtium had its full genome characterised based on RNA extraction and HTS of mashua samples: opium poppy mosaic associated virus, OPMaV, (see subsection 4.2.4). Mashua may serve as a reservoir of viruses, introduced or indigenous, which may later be transmitted garden nasturtiums growing in the UK and the EU.

Species origin In some cases constructing a phylogeny of a virus based on sequence information can help suggest a plausible a centre of origin for its species, even if surveys did not. This approach allowed the hypothesis of a South American origin for TropPV3, and possibly the putative novel virus from which a fragment was sequenced (Figure 4.2.5). Regardless of whether the 1kb fragment is a novel virus or a divergent MasVY isolate, these viruses are both likely members of the potato virus Y subgroup. Though potyvirids are likely to be polyphyletic (Gibbs et al., 2020), this subgroup is thought to originate with wild and cultivated solanaceous crops in Bolivia and Peru (Spetz et al., 2003). For this reason, it is likely that these two potyvirus(es) detected in mashua are South American in origin, and came to Europe with the ancestors of the plant tested.
Of interest also is the difference in the observed virome of European and non-European plants. Including the putative novel viruses described during the present work, considerably more

viruses have now been reported in the post-HTS virological era than have been described in ARTCs in the entire history of study of viruses in ARTCs (see Table 1.2 on Page 20).

Relation to pathway risk It appears that ARTCs are accumulating at least some European-indigenous viruses during their stay on the continent. Inversely, some viruses were detected in ARTCs sampled in South America but were not detected infecting ARTCs in Europe: for example, PVT and PBRSV were not found in any European plants in the literature or during the present study. PYV however, was detected in yacón purchased from European sellers and from the commercial nursery in Ireland. *Potato yellowing virus* as a species probably has a South American origin (Silvestre et al., 2020).
Some properties of the internet trade pathway itself are also worth consideration, as they present their own challenges relative to the conventional paradigm of large shipments by dedicated traders. In particular, some of the trade in ARTCs in Europe is illegal, including some of the purchases from this study which lacked correct customs documentation and plant passports. This lack of documentation makes tracing an outbreak difficult even if it is detected.
For example, an English hobbyist grower may purchase oca tubers from a supplier in the UK. These tubers may be grown in the seller's allotment or garden and sold from there. They may be purchased by the seller from another seller, perhaps from another country, e.g. from drop shipping. Drop shipping has caused difficulties in epidemiology before, for example during a US outbreak of salmonellosis linked to back-yard rearing of mail-order poultry in 2013 (Anderson et al., 2016). Even if the plants for sale are grown by the seller, the propagules that led to these plants may have been themselves imported from other countries in Europe or further abroad in the recent past, either in the previous generation, or further beyond. This is an additional risk to be accounted for when investigating the trade of a primarily clonal crop.
This lack of information provided by sellers, or in some cases intentionalobfuscation of the truth, will muddy any attempts to determine the source of an outbreak. This is especially important considering that an infected collection of propagules may have already found its way into any number of back gardens by the time an infection is detected in a shipment (see chapter 1).

7.5. 2.2 Assessment of the Probability of Introduction and Spread

7.5.1. 2.2.1 Probability of Entry of a Pest

Entry risk can be expressed as the number of 'affected pathway units' per unit of time, e.g. per year (Jeger et al., 2018). Probability of entry is twofold under a propagule pressure model: one is the rate of arrival of the commodity, the other is the probability of the commodity being associated with 'the pest' (e.g. interception rate at port) (Simberloff, 2009). The two together form the rate of arrival, which is why trade volumes may be used as part of the assessment of risk, especially useful as it provides immediately actionable links in the import chain once identified (Burgman et al., 2013). In the case of e-commerce, rate of arrival may be difficult to determine due being a dynamic trade with a large number of small sellers (Humair et al., 2015).

High throughput sequencing of viruses extracted from plants available via this pathway allows a rough assessment of probability of association of pest(s) with a commodity. That is to say: if I purchase oca plants from eBay, how likely is an individual purchase to contain viruses of some kind in at least one plant? PCR may then resolve this into individual plants (percentage infected) and determine rates of mixed infections. The answer, as seen in this thesis, is 'highly likely' in Europe. Any given purchase of oca tubers is highly likely to contain tubers infected with at one virus. RNA was extracted and sequenced from leaf tissue grown from the tubers during the present study, indicating that the viruses detected infect daughter plants from infected tubers at least some of the time.

Tubers surveyed in this thesis The mashua tuber planted for the qRT-PCR survey, once planted, were determined to be free of the oca viruses but exhibited symptoms associated with TropMV or another virus, subsequently dying. Without a test, it can't be assumed that the symptoms were viral, as mashua may respond to abiotic stressors from transport this way.
A small number of the oca plants purchased during this study were free of the viruses tested via qRT-PCR, but only two batches were unambiguously free of all of the four viruses tested for, the 'Cream' batch and the 'Oca de Perou' batch (subsection 3.2.15). These plants could have been infected by, for example, ArMV or UlPolV1, as these viruses wasn't tested for using qRT-PCR. The oca plants which tested positive for ArMV using primers designed during this study (subsection 3.2.5) did not test positive when assayed using ELISA with commercial ArMV antiserum from DSMZ. This is likely due to substantial differences in the coat protein between the DSMZ biotype and the oca biotype, which is the part of the virus (the antigen) that ELISA as a method detects through binding to the assay-specific antibody (Hull, 2021). Even if these oca plants were tested for ArMV via ELISA as part of border control during an ordinary purchase, the ArMV infection likely would have evaded detection were it not for HTS. This is despite the ability to transmit the virus mechanically from infected plants to healthy *Chenopodium quinoa*, confirmed by symptom expression and, more importantly, PCR. Nucleotides of the ArMV population infecting the oca plants sampled during this study was also not amplified using the broad-spectrum ArMV primers (which were designed at Fera specifically due to poor detection of some ArMV biotypes with existing prior primers). These primers intentionally used a variety of isolates from around the world in their creation, with aim to detect as many biotypes as possible. This high variability of viruses on the molecular level is a known difficulty when detecting them using ELISA and PCR (e.g. in grapevine (Velasco and Padilla, 2021)). New variants of known viruses which escape detection with current methods are still being discovered, like the oca ArMV biotype(s). HTS was able to detect the ArMV sequence in the infected oca plants where ELISA did not give a positive result.

E-commerce risks In terms of the other two qualities of e-commerce, which are the speed of transport related to the survival of the pest, and the likelihood of surviving existing pest management scenarios, oca tubers demonstrably retain infectious virus for long enough to be purchased and planted. They are evading phytosanitary inspections at the border. This is evident

both in the observed mislabelling and in the English plant that crossed the England-France border at least once before being purchased under license for this study. Viruses can't be managed if they aren't known to be present, so one may consider detection fundamental (Fontdevila Pareta et al., 2023). It is essential to develop a picture of what viruses are already present in the UK and to obtain a snapshot of viral 'stowaways' being traded, and this latter portion is made very difficult by small, fast-traced courier parcels arriving stochastically in back gardens around the country (Giltrap et al., 2009).

Complications of unknown sample origins Little is known about the origins of the ARTC propagules circulating on the European internet market today. As mentioned here, the oca germplasm that led to the Aoteoroa-New Zealand oca lines and at least some of the European lines seems to originate from market purchases in South America (Fletcher et al., 2005). This was prior to the Nagoya protocol, which was adopted in 2010 and came into force in 2014 (Secretariat of the Convention on Biological Diversity, 2011, Schindel et al., 2015). Some of the oca lines in this study clearly derive from Aoteoroa-New Zealand ('Halford's Red') (Figure 3.2.1 on Page 53).
Other tubers were labelled with more South American names, such as 'New from Peru', but this is less demonstrative, as every oca plant has an ancestor somewhere in Peru, though perhaps not recently. Still others have no indicator of where the original germplasm came from that began the European branch of their clonal line ('Long Yellow' for example, or 'Cream'). Only one oca plant developed tubers during the study, the 'Giggles' plant which was purchased with tubers forming in the pot, so it is difficult to say for certain which individuals are adapted to a Northern European day length. Such an adaptation might indicate a recent origin in Europe or Aoteoroa-New Zealand due to selective breeding efforts. Perhaps a phylogenetic study of oca sourced from multiple sources, including the Andes, European countries and Aoteoroa-New Zealand, could resolve clusters and help elucidate the provenance of these 'varieties' relative to one another.

7.5.2. 2.2.2 Probability of Establishment

Once a pathogen is understood to be associated with a pathway, the likelihood of its establishment after entry must be ascertained as part of understanding the risk posed. Presence of a virus does not automatically cause an outbreak (Vurro et al., 2010). The present work used two means to ascertain this:

1. Investigation of close relatives for novel viruses known to be associated with oca (i.e. potential inoculum sources should an outbreak occur)

2. Mechanical inoculation attempts to investigate host range. These form some early explorations of the potential epidemiology of these viruses, which must be understood in order to know these risks and protect plant health.

For example, as noted above, presence within UK members of *Oxalis spp*. suggests that OCV1 is already present in Europe, making it plausibly less of a threat. It is less likely that the findings indicate that this virus was introduced and has already become widespread, though that is not impossible. The finding of OCV1 in *Oxalis sp.* growing in Poole, Dorset also suggests that what is quite an abundant plant within the UK can share at least one virus with oca. *Oxalis* species members are common ornamentals (Young, 1958)(and OCV1 was also detected in *Oxalis* 'Irish Mist' (subsection 3.2.15). embers of *Oxalis sp.* are frequently part of English woodland and hedgerow flora; they can grow as weeds; and they grow as volunteers in cracks in the pavement in public places or private land (the source of the 'feral' plants sampled for this study)(Young, 1958). Given that the ulluco plants found to be infected with quarantine viruses in 2019 originate from an allotment in the South of England (Fox et al., 2019), it seems plausible that oca plants would be grown in the same circumstances.

If a virus is introduced on oca, a chain of transmission might develop in the UK. For example a virus might spread from oca to *Oxalis spp.* that could be growing in cracks in the pavement. These could be vectored by nematodes, aphids, or humans. Perhaps aphids may convey the virus to oca or other *Oxalis sp.* in neighbouring allotments. While South African *Oxalis spp*. are all geophytes, South American *Oxalis spp.* display a range of growing habits and life strategies i.e. they fulfill different ecological niches (Dreyer et al., 2006). This could well lead to an increased availability of a 'suitable host' during different seasons, for example, potentially leading to temporal continuity of inoculum sources throughout the year. Some of the oca variants circulating in Europe are likely to be adapted to the European day length and thus may be grown over several vegetative generations, accumulating virus with time, leading to potential spillover and spillback.

The presence of OCV1 in only a few random spots in the UK could be due to the presence of a suitable vector, seed transmission or, as discussed, possibly an artefact of an ancient integration into the germline. Moreover, OCV1 was never associated with symptoms, let alone casually. It seems reasonable that OCV1 specifically isn't a particular threat to UK and EU plant health. The survey allowed elimination of one candidate from consideration, but it also highlighted a potential epidemiological network that could emerge in future, by demonstrating an overlap between the virome of an imported crop and established wild-growing flora.

The uncertainty of species identity of some of the purchased ornamental plants does highlight a wider issue of incorrect or ambiguous labelling of species or genotype, intentional or accidental, of plants from the internet and even some commercial retailers (Van Den Neucker and Scheers, 2022, Pinczinger et al., 2020). This could make assessments of a virus' host range a challenge. For example, the plants labelled as *O. triangularis* may belong to this species or another species in the genus. In the plant world cryptic species exist which can result from sympatric speciation or allopatric speciation and may be distinguishable only by genomic information (Govindarajulu et al., 2011). Building a potential viral distribution range based on host distribution thus becomes slightly more complex when considering members of a new plant's genus as sources of viral inoculum.

7.5.3. 2.2.3 Probability of Spread after Establishment

Spread after establishment is the final phase of invasion as outlined by ISPM 2, 11 and 21. It is the most difficult to quantify as it factors in geography, physical barriers and vectors. Models are further complicated by long distance jumps of hosts within the area under assessment due to human movement of material (Robinet et al., 2012).

Seed sharing initiatives, such as the Guild of Oca Breeders, led to oca plants being swapped between English counties (section 1.6 on Page 12), whilst the online trade provides a regular availability of tubers, especially after the harvest in winter. Although the material is meant for planting, it may not be regulated as such when they are declared on customs as 'flavourings' or 'cosmetics', leading to unregulated trade.

In the case of ocas, movement due to vectors is likely to play a part in the potential spread of the viruses infecting oca plants. A number of the viruses described from sequence detected in oca in the present work are related to viruses which are known to be spread by aphids: capulaviruses (Ryckebusch et al., 2022, Varsani et al., 2017), poleroviruses (Hoffman et al., 2001), potyviruses, caulimoviruses, cytorhabdoviruses and nucleorhabdoviruses (Ng and Perry, 2004). With nematode-transmitted viruses, like some nepoviruses, movement of planting material is a major pathway for both the virus and their nematode vectors, as was the case for GFLV (Raski and Hewitt, 1963).

In the native Andes, PVY is most common at less than 3,000 masl due to its aphid vector's poor survival rate at cold temperatures. In contrast, the mechanically transmitted transmitted potato virus S and potato virus X are found at relatively higher elevations in potato (Fuentes et al., 2019). In the Andes this phenomenon provides a strategy for virus management, moving potatoes up and down elevations to take advantage of physical barriers and temperatures that interfere with the movement of aphid vectors of viruses (Thiele, 1999).

This is interesting to view in maca, which is grown at high altitudes (>3000 masl) in Peru and in China but sometimes in China also at lower altitudes of 2,400-2,500 masl. Cultivation of maca in Yunnan, China was greatly increased between the 2005 and 2015, though has more recently fallen into decline (Yin et al., 2019). Maca from Yunnan, China was fairly frequently infected by aphid-borne potyviruses and poleroviruses according to investigation of SRA datasets derived from transcriptomic sequencing of maca (section 6.2 on Page 138). Whether these viruses are present in Peruvian maca plants that have yet to be sequenced and detected remains to be seen, but given the frigid temperatures of the Puna, perhaps not. Indeed a plant grown at altitudes where most insect vectors would struggle to survive may well not have been under selective pressure, by nature or by humans, to develop a particular resistance to these diseases. Sampling of maca from the Puna would provide a better insight on this.

Identification of a virus, and determination of its relation to other viruses, can help in risk assessment, but with some caveats (Fontdevila Pareta et al., 2023). Mashua is host to a number of potyviruses, some of which were first described during the present work (TropPV3 and a possibly second novel potyvirus). OVX, a potexvirus, was also described from sequence derived from oca samples during the present work. Zucchini yellow mosaic virus, a contact-transmitted

potyvirus, can survive for six hours on clothing and footwear and then be able to infect plants (Coutts et al., 2013). Potexviruses such as PVX and pepino mosaic virus are known to be easily transmitted by mechanical contact. Some mechnically transmitted viruses spread via contaminated clothing and tools, with the tobamovirus, tobacco mosaic virus able to infect new plants from soil for as long as 20 months after the infected plant host is removed (Sastry, 2013b). This method of spread would require that TropPV3 and OVX share both this contact-transmission trait with members of their genus and that a suitable host is available, but it isn't beyond reason that mechanical transmission could be a potential pathway of spread for these viruses after entering the UK in infected ARTCs.
Extrapolating in this manner must be done with caution (Fontdevila Pareta et al., 2023), but it provides a way of formulating hypotheses for later testing. On identification of a nepovirus, one might wish to survey for nematodes of the corresponding vectoring taxon in the new region, for example. Such information may already exist and be available via a literature search (Jeger et al., 2018).

7.6. 2.3 Assessment of Economic Consequences

Economic consequences can be a direct result of pest damage or indirect due to trade restrictions emerging from phytosanitary action (IPPC Secretariat, 2013). Direct effects are in theory easier to determine than indirect ones (which are the result of human actions), but in practice complex. In some cases the viruses detected infecting ARTCs collected during this study have known economic consequences. This is the case also for some of the viruses which were detected by assembling reads from the HTS datasets available on the SRA which were derived from transcriptome sequencing of ARTC samples. These viruses which are well-characterised include ArMV, sugarcane mosaic virus, TuMV and PYY. For such well-understood viruses an assessment of the risk posed to European plant health can be made more easily as information is already available on their incidence, symptoms and vectors. For the novel viruses, lack of causal association with symptoms may mean that they are not a plant health risk, but this is not certain. In terms of environmental consequences, emerging plant pathogens that have evaded human control efforts already threaten wild ecosystems. *Phytophthora ramorum*, for example, infects trees such as oaks, endangering plant communities with cultural and amenity value (Green et al., 2021). This is especially troubling in cases where a virus and host are newly introduced to one another, which may cause greater disease severity as the two have not had time to co-adapt (Jones, 2009).

Yácon Of the A1 pathogens listed above infecting ARTCs, only PYV was detected in any of the European plants (yacón). The propagation of yacón in the UK was already subject to a warning sent out by DEFRA in 2022; its import is currently prohibited from all third countries (Barker, 2022, DEFRA, 2024). Apart from assembly of a small sequence with homology to garlic virus A and a small sequence with homology to cucumber green mottle virus, which may originate from a contaminant, all the viruses detected in this thesis were already known to infect

yacón in the relevant literature. The data here support the notion that PYV is widespread in yacón and that importing them constitutes an unacceptable risk of this virus spreading to UK potatoes. However, it should be noted that, at very low incidence, PYV is unlikely to be a virus of concern to potato in Perú where the host and plant are co-adapted (Silvestre et al., 2020).

Ulluco For ulluco, likewise prohibited for import in the UK from non-EU countries (DEFRA, 2024), the virome of the plants provided by the commercial supplier in Ireland was similar to that of the plants found in a planting in the South of England, presumably derived from internet-traded stock (Fox et al., 2019). Notably the tubers from this thesis were infected with the following viruses which were also detected in the 2019 study of viruses in ulluco from a UK allotment: UlPolV1, Ullucus tobamovirus 1, Ullucus comovirus 1, papaya mosaic virus (ulluco strain, PaPMV-U) and BBWV2. UlPolV1, Ullucus tobamovirus 1 and Ullucus comovirus 1 have not been associated with any symptoms and were first reported in the UK. It is not impossible Ullucus comovirus 1 and Ullucus C comovirus are two names for the same virus (Fox et al., 2019).
Broad bean wilt virus 2 has a broad host range including lettuce and petunia, and wide geographic distribution. BBWW2 viruses can cause economically important damage on a variety of crops and ornamental species (Ferrer et al., 2011).
PapMV-U was first described from an infection in ulluco plants in 1982. The PapMV strain described infecting ulluco is biologically different from other strains (Brunt et al., 1982b). In mixed infections with Ullucus mosaic potyvirus (UMV) and Ullucus C comovirus, PapMV-U reduced ulluco yields by as much as 38% (Lizárraga et al., 2001). It has not been shown whether this same yield reduction might apply should this virus complex infect other crops. In addition, ulluco grown from virus-free seed developed a low incidence of infection with Andean potato latent virus, UMV, Ullucus C comovirus and PLRV after just one field generation (Lizárraga et al. 2001).
Tomato spotted wilt virus (TSWV) has been detected in European ulluco samples (De Jonghe et al., 2022). TSWV has an extremely broad host range, is widespread, and is economically damaging, including on potato (Abad et al., 2005). Inection by TSWV is one of the major constraints on tomato production (Gupta et al., 2018).

Mashua Mashua virus Y, at least one novel potyvirus, and at least one novel polerovirus (novel in sequence; they may be already characterised) were all detected in mashua plants during the present study. These detections came from samples purchased, extracted and sequenced with intent to detect viruses and from assembling and investigating sequences from the SRA. Though sequences with homology to other viruses were assembled from the reads contained in the SRA datasets, the sequences have too low a coverage of the viruses in question to diagnose the plants sampled with any confidence. In addition, the reads were sequenced from nucleotides extracted from a Bolivian sample taken decades ago. The virus may not still be widespread in Bolivia in the present day, and even if it is, this does not mean that current-day European mashua plants are

infected with the same viruses as were detected in the datasets derived from sequencing of extracts from the older mashua samples.
The detection of these viruses in the sequence data could indicate that mashua is a potential host for them, but may not necessarily suggest that they are circulating with mashua and that infected mashua plants are likely to cause outbreaks of these viruses. The findings from the Phytosanitary Risks of Newly Introduced Crops (PRONC) project which preceded this thesis would suggest that these other potyviruses (wild potato mosaic virus, potato virus V and papaya ringspot virus) and chrysanthemum stem necrosis virus are not widespread in European mashua (De Jonghe et al., 2022).
Two carlaviruses have been reported infecting mashua in the literature, and one was detected from data-mining the one SRA dataset derived from sequencing of RNA extracted from a mashua plant (section 6.2). These are red clover vein mosaic virus and potato virus S. Red clover vein mosaic virus causes severe disease and even death in leguminous crops such as peas, chickpeas and lentils (Larsen et al., 2009). Potato virus S was detected only from assembly of reads from the SRA dataset that represented a Bolivian mashua sample. The sequence obtained in this way was short in length with low coverage of the potato virus S genome, possibly indicating that the detection was due to contamination. Potato virus S was first isolated in the Netherlands and reduces yield of potatoes, with economic implications for seed potato; it is likely to be contact transmissible (de Bokx et al., 1972).
TuMV was detected in *T. majus* in Aoteoroa-New Zealand in a mixed infection with opium poppy mosaic virus, broad bean wilt virus 1, and verbena latent virus (Ochoa-Corona et al., 2010). Broad bean wilt virus 1 was found in UK ulluco plantings before (Fox et al., 2019). Possibly, mashua could be infected with any of the viruses known to infect the closely related *T. majus*, which could expand the number of viruses known to infect more than one ARTC. TuMV was detected in maca plants from China in 2015 (Yin et al., 2015) and from assembly of some of the SRA datasets derived from sequencing of extracts from maca plants (section 6.2), providing another shared virus between two ARTCs if mashua were to become infected with TuMV, as *T. majus* can be.
Lastly Physostegia chlorotic mottle virus has been detected in mashua from Europe (De Jonghe et al., 2022). Physostegia chlorotic mottle is widespread in Europe and has been implicated in disease of tomato, aubergine and cucumber, including ripening abnormalities in the fruit (Temple et al., 2022). Mashua tubers are thus capable of carrying a number of economically important diseases over long distances, remaining undetected.

Maca Maca was not available for sale on European websites at the time of this study, searching with the name maca or with either the correct (*L. meyenii*) or incorrect (*L. peruvianum*) latin binomials (see (Hermann and Bernet, 2009)). Nonetheless, the detection of sequences representing TuMV, sugarcane mosaic virus and turnip yellows virus in SRA datasets that originate with transcriptome studies of maca plants suggests that maca, along with other ARTCs, could act as a reservoir for economically important viruses if grown in the UK in the future. This frequent detection may indicate that the vectors of these viruses are present and

feeding on the maca plants which constituted the biological material for the relevant BioSamples. Such vectors likely feed both on maca and on other plant species which are hosts of some of these viruses, such as maize and oilseed rape (Wu et al., 2012; Filardo et al., 2021). For example, TuMV has a very wide array of vector species, making it easier for the virus to spread than viruses which may have a limited range of vectors available (Walsh and Jenner, 2002). This presence of the viruses detected in maca-derived datasets, and their respective vectors, within the UK could lead to an outbreak of these introduced biotypes.

Oca Arabis mosaic virus has been detected in oca before, and this was noted as a risk that internet trade in oca plants might pose to plant health in Europe (De Jonghe et al., 2022). Arabis mosaic virus is mostly economically damaging when it infects grapevine, and was predicted to have only "unlikely to moderately likely" risk of spread via plants for planting in 2013, with a very likely chance of establishment and a likely spread via natural means thereafter (EFSA Panel on Plant Health, 2013). BRSV was detected in oca plants provided for this study by the commercial supplier in Ireland, though at a low confidence that this detection represented a genuine infection. While BRSV is not presently regulated in the EU (EPPO, 2023a, EPPO, 2023b), it is biologically similar to the regulated non-quarantine pest, tomato black ring virus (Fowkes et al., 2021a).

Generally there is a distinction of three main 'groups' of viruses which infect oca, based on geographic information. The first is viruses known to infect oca plants in the Andean region from older literature. The second is viruses detected only in oca plants sampled in Europe, and the third is viruses that may be South American in origin and were detected in samples of oca grown in Europe. In the case of the first, such as AVB, virus elimination work was undertaken with the express purpose of removing those viruses from oca when imported from South America to Aoteoroa-New Zealand, as detected by (Fletcher, 2001). These viruses which were detected and eliminated from oca in Aoteoroa-New Zealand are notably absent from European samples.

In the second case, viruses so far detected only in oca from Europe could have entered the oca population at any point either in South America or elsewhere. ArMV, for example, has only been reported in the literature in Europe, and is a virus that is European in origin (Je Jonghe et al., 2022; Digiaro et al., 2017). It is thus reasonable to assume that ArMV has emerged onto oca from European flora.

The third group, plausibly Andean viruses that were not reported before the widespread implementation of HTS, *are* sometimes detected in oca samples purchased from sellers in Europe. This includes plants which may have been grown in the UK, notably the detection of UlPolV1. Taken together, this does lend credence, along with the day-length adaptation, to the theory that the oca tubers presently circulating online in Europe derive substantially from commercial lines from Aoteroroa-New Zealand. Specifically, oca in Europe may descend chiefly from lines which were subjected to virus elimination in the past, but only checked for viruses which were then understood to infect oca (PapMV, AVB and UMMV) (Fletcher, 2001). PapMV, AVB and UMMV were not detected during this study in oca plants from Europe, but the viruses

from the Andean region which were not subject to elimination work may have 'slipped through' and later found their way to Europe in infected oca plants.

Clean oca plants The oca plants from the commercial nursery in Ireland were largely determined to be free of virus, all except for tubers of the 'Halford's Red' variety, at least one of which may have been infected with BRSV. This is good evidence that a grower with phytosanitary interest maybe able to provide quality seed that presents minimal plant health risk. Such material may be a good candidate for plant health certification,either formally or informally. However, this certification process is not free. Such a certified product is likely to come with a higher price tag to offset the increased production costs (Jones, 2006). Given the general lack of knowledge around plant health among European small growers and civilians (Michi et al., 2023), cheaply available, poor quality, frequently virus-infected plants are going to remain an attractive option for many growers. A survey of UK buyers found that consumers tended to have poor awareness of biosecurity and expressed concern that accredited products would be more expensive even when stating support for such a scheme in theory (Dunn et al., 2020).
While a grower might discard obviously symptomatic plants, these viruses seem to be associated with minimal, if not no, symptoms and management by removing symptomatic plants is thus not likely to be effective. While this lack of associated symptoms is favourable news as it means that most of the viruses infecting oca are unlikely to be especially damaging, viruses are biological entities and they do, of course, evolve. Given the rate of mixed infections in both the informally traded and commercially available plants, there is risk of a damaging recombinant emerging. However, long-term vertical transmission of all parasites, including plant viruses, alters the selection pressures, favouring lower virulence (as there is no need for high accumulation to increase the chances of horizontal transmission) and survival of the host to reproduce (Pagán et al., 2014).
Even cryptic viruses may, in rare cases, be transmitted horizontally, making the boundaries between lifestyles permeable; though lack of viral diversity in truly persistent-lifestyle viruses may be a barrier to developing or re-developing an acute lifestyle, cryptic viruses may contribute to damaging viral variants via recombination (Roossinck, 2010).
In short, the presence viruses with these niche crops which may or may not affect host fitness may represent a multiplicative rather than additive risk. This is the case for both the putative novel viruses described in this thesis and infections with viruses known from the literature, such as ArMV, which was apparently asymptomatic on the oca plants but is damaging on grapevine (Digiaro et al., 2017).
To use an arbitrary example: if virus A reduces yield by 20%, and virus B by only 2%, and have no synergistic effects for a combined reduction of 22%, there is still a risk of virus C emerging through recombination of the two. Virus C may be more damaging (e.g. intraspecific recombination in PVY strains (Boonham et al., 2002)), expand the host range of virus C relative to its parents (e.g. interspecific recombination in potyviruses leading to watermelon mosaic virus (Desbiez and Lecoq, 2004, Moury and Desbiez, 2020)), or allow resistance breaking (e.g. intrafamilial recombination between members of the families *Luteoviridae* and *Tombusviridae*

(Miras et al., 2014)). Such a recombinant, in this arbitrary example, might reduce the yield by some 40% instead. Recombination is likely to be responsible for the ability of PYV to infect potato, which also infects yacón and pepino (Silvestre et al., 2020). This recombination risk is broadly covered under the 'adaptability' portion of the biotic factors of ISPM 11 (IPPC Secretariat, 2013).
Due to the poor condition in which some of the tubers arrived, it was not possible to associate symptom development with viral infection, as any signs of ill health could be due to tuber storage or some other biotic or abiotic cause. Symptoms were especially difficult to determine as oca is not a crop whose symptom expression is well documented.
Lastly, the lack of associated symptoms for these viruses does not mean that they are a) truly asymptomatic (as noted, the plants presented with various symptoms which could not be linked to any particular cause) or b) that if visibly asymptomatic, they do not affect yield or vigour.
In sum, oca is likely to represent the least threat to European crops and wild plants out of the ARTCs studied. Yacón and ulluco have already been banned in the UK due to unacceptable levels of risk, and perhaps mashua ought to join them, host as it is to several economically important viruses of other crops. Maca is not presently grown widely in Europe but should be investigated in case of continuing interest in niche crops, as it is host to a number of aphid-vectored viruses and may act as an inoculum source for infections on economically important crops.

Viruses shared with potato Viruses have been shared between different members the ARTC group before, and between ARTCs and potato. Though the PLRV isolate once reported in ulluco is likely to be its own species, as is the APLV isolate (Lizárraga et al., 1996a, Fox et al., 2019), some other viruses may infect more than one ARTC species and/or potato. Examples include UlPolV1 in both ulluco and oca (this study), potato black ringspot virus in oca, arracacha and potato (Jones, 1981, Lizárraga et al., 1994), and the aphid-transmitted potato yellowing virus (PYV) in potato and in yacón, though they group separately (Silvestre et al., 2020). As seen, this shared virus range may extend to turnip yellows virus in both maca and mashua as well, and plausibly broad bean wilt virus 1 could infect mashua as well as ulluco, given its reported infection in *T. majus* (Ochoa-Corona et al., 2010).

7.7. Conclusion of PRA

Two considerations of the IPPC stand out when studying the data here: 'cooperation in the provision of information' and 'avoidance of undue delay'. Cooperation in the provision of information might include data sharing, including access to information on public archive, such as the SRA datasets utilised in this thesis. It also has associated dangers, such as concerns about mining of SRA data to justify a discriminatory trade practice, or effects on trade due to potential presence of a virus known only from sequence data (Macdiarmid et al., 2013, Fontdevila Pareta et al., 2023). Nonetheless, re-use of data and cooperation between parties is essential if the sheer

volume of information available to virologists and risk assessors is ever to be managed (Fontdevila Pareta et al., 2023, Kutnjak et al., 2021, Olmos et al., 2018, Hou et al., 2020). When avoiding undue delay, biological characterisation continues to be thorny. It is difficult to categorise viruses for several reasons:

1. They vary greatly in biology and each must be studied in a bespoke manner fitting the characteristics of the entity being studied.

2. In some cases, viruses are so divergent from known viruses that they defy understanding and it can be troublesome to know where to begin, as evidenced by the number of viruses belonging to new genera and even families known only by sequence (Simmonds et al., 2017).

The large amount of data available on viruses studied *because* they were easily categorised may mean that a number of viruses that are more difficult to manipulate experimentally (e.g. isolate, inoculate and maintain) may have been overlooked in the past. It is not possible to say with any certainty if these viruses, whose qualities make them hard to study, are less, equally or more abundant than viruses which are amenable to mechanical inoculation and other classical techniques. The difficulties with inoculating the novel nepovirus experimentally onto indicator plants demonstrates that it can take many working hours to elucidate even the smallest facet of one virus' biology. Depending on how one defines 'undue delay', full characterisation of every organism discovered through HTS may not harmonise with IPPC directives at all, especially as HTS is more frequently utilised for front-line diagnostics, detecting non-target novel viruses (Maree et al., 2018).
Despite the challenges, the properties that make a virus easy to study and the properties which make a virus a noteworthy pest can align in some ways. For example, evidence of establishment in other areas can be taken as an indication that an organism is likely to be a pest (IPPC Secretariat, 2019). Although poleroviruses are not considered mechanically transmissible on the whole without biolistics (Hoffmann et al., 2001), viruses with a broad host range are readily mechanically transmissible and considered easily spread, e.g. Arabis mosaic virus (see chapter 5). Different methods have different sensitivities, but generally the detection of a virus by any method increases with concentration in the cell tissue being studied i.e. the titre (Villamor et al., 2019). Titre has in some studies been found to be correlated with symptom severity and yield loss in different viruses (Van den Bosch et al., 2006a, Zhu et al., 2010). However, phloem-limited viruses such as luteovirids, which are an economically important group (Hoffman et al., 2001), tend to be found at low titres, complicating analysis (Kawchuk et al., 2002). The chapter on the host range of nepoviruses (chapter 5) outlines the dangers of assuming risk of a novel virus based on other members of its genus, especially because certain genus members may be 'over-studied' based on the ease of performing experiments with them. To dismiss a novel viruses as economically unimportant merely because it is difficult to characterise would have been a mistake. However, the encountered difficulty, in the present study, to easily transmit some of the viruses described in this present work with known

experimental methods may mean that they are not a high priority in terms of risk. Meta-analyses of studies of viruses discovered via HTS in fruit trees, tomato and *Poaceae* indicate that time- and cost-effective methods tend to be better represented in the literature (Hou et al., 2020, Rivarez et al., 2021, Fontdevila Pareta et al., 2023). Indeed, an updated 2023 framework providing suggestions for dealing with viruses detected via HTS acknowledges this and suggests finding ways to leverage and improve these cheaper, faster methods for risk assessment (Fontdevila Pareta et al., 2023). The same resources applied earlier and thus more effectively is easier to 'sell' to policy makers than simply asking them to allocate more resources (Ward, 2016). Efforts to understand a virus and perform PRAs only after severe outbreaks have occurred leads to late action and poorer phytosanitary outcomes (Brasier's 'stable door' syndrome) (Brasier, 2005, Ward, 2016). This is a limitation of the present PRA model, which focuses on individual organisms rather than on pathways (Brasier, 2005, Evans et al., 2010). Prioritising studies based on potential risk becomes more important with time due to increasing global plant trade, movement of people and plant material and ongoing economic consequences of wars, which affect management of pest invasion risk in Europe (Michi et al., 2023).
Perhaps the most troubling finding from this thesis, however, is the demonstrated illegal plant trade across borders. This requires more study to continue to determine the true risk that this illegal trade is posing to plant health in Europe, and how it can be managed. The virome of plants which could easily have entered the country for the purposes of planting has been outlined here, and it is possible that other plants, possibly hosts of other quarantine pathogens, are evading border inspection as well. Because of the increasing rate of e-commerce and deliveries of plants via small parcels, tighter measures are needed to prevent a severe biosecurity breach in the near future.

7.8. Speciation and Divergence

As can be seen, some viruses detected in ARTCs infect more than one ARTC host. Some ARTC species are known to be hosts of several viruses which may be closely related (serologically or by phylogenetic comparison of sequence). Sometimes these closely related virus may infect the same ARTC individual in a mixed infection, e.g. the mashua plant which was found to be infected with two closely related poleroviruses. These two poleroviruses cluster separately from both oca and ulluco isolates of UlPolV1 and away from cereal yellow dwarf virus and pPLRV as well. Mashua virus Y and the novel potyvirus reported here are both PVY subgroup viruses which infect mashua, and both viruses probably originate in the Andes (Fuentes et al., 2019; Adams et al., 2018c).
This speciation of some of the viruses detected in ARTCs might be a case of allopatric or parapatric speciation caused by the diverse geography of their native mountain range. This physical segregation due to variations topography and elevation has led to speciation of various flora, contributing to exceptional biodiversity in the region (Vargas et al., 2023, Smith et al., 2014). Human movement of germplasm would make the boundaries porous, but different landraces are adapted to different conditions, and in some cases farmers replace their seed very

infrequently (Thiele, 1999). In other words, the cluster of closely related but distinct virus species found in the different ARTCs may be due to some sub-populations of ARTCs remaining distinct for long periods of time.
The fact that different viruses are now being found together could be the result of differing genotypes of mashua, oca and ulluco being brought together at the same (low) elevation in the same place by growers in Europe and Aoteoroa-New Zealand, presenting more opportunities for cross-host spread. This may not have been as common in the agricultural practices in the Andean region.
In plant-virus systems, human activities can influence the broader strokes of viral evolution. For example, the movement of potato via trade is thought to have shaped the genome of PVY thanks to the relatively recent and rapid expansion of its geographical range, favouring less intense negative selection than TuMV (Gibbs et al., 2020). Disease control efforts also shape viral populations. This can be direct in the form of the 'arms race' to escape host resistance (resistance-breaking) and then needing to breed new resistant varieties (Kobayashi et al., 2014, Walsh and Jenner, 2002). It can also take the form of inadvertently selecting for damaging traits in a viral population, such as a model which found that some virus control efforts in vegetative crops may select for viruses which accumulate to a high titre (van den Bosch et al., 2006b). Other changes can can be due to germplasm exchange during breeding efforts. PVY isolates were in some cases found to be likely 'remigrants', whereby European (including UK) PVY found its way back to the Andes and Chile when introduced for breeding programs before extensive viral testing (Fuentes et al., 2019).

7.9. Considerations of Throughput Sequencing

The present study used total RNA with ribosomal depletion to enrich for viral sequences. Different enrichment methods have different strengths and weaknesses and should be selected according to the desired outcome. A comparison between dsRNA and ribosome-depleted total RNA extraction found that HTS after totRNA extraction produced higher coverage of high-titre viruses but that HTS after dsRNA extraction was better able to detect viruses present in very low concentrations, including persistent viruses (García et al., 2020). This concurs with the findings of a comparison between datasets derived from dsRNA, siRNA and totRNA extraction methods, which found that sequencing of nucleic acids extracted with the dsRNA method detected all viruses in the dataset, sequencing of extracts obtained via the totRNA method did not detect Vicia cryptic virus and sequencing of extracts obtained by the siRNA method did not detect the novel Wuhan aphid virus 2 (Gaafar and Ziebell, 2020). Sequencing of extracts obtained via the totRNA extraction method may fail more often to detect low-titre viruses (such as cryptic or persistent viruses) without sufficient depth as compared with other extraction methods, a drawback noted by a review of HTS methods in 2015 (Roossinck et al., 2015).
Another study investigated viruses with different genome organisations by sequencing extracts obtained by the totRNA and siRNA methods, finding that coverage differed between Baltimore classifications for each method: sequencing of circular ssDNA viruses and viroids had higher

recovery with siRNA extraction than tRNA extraction, but recovery was higher using tRNA extraction for most linear ssRNA viruses tested (Pecman et al., 2017).
Sequencing nucleic acids extracted with VANA with a 454 sequencer led to longer assembled contigs than with extractions of siRNA on an Illumina sequencer, though the two together were needed to assemble the complete genome of a novel mastrevirus (Candresse et al., 2014). An analysis of wild plant viromes using nucleic acids extracted using the dsRNA and VANA methods found that dsRNA extraction outperformed VANA extraction in the number of reads mapped to contigs, number of virus families detected, and number of operational taxonomic units recovered (viral richness), but detected more DNA viruses when nucleic acids were extracted with the VANA method (Ma et al., 2019). A later benchmarking study in a mock virus community likewise found that virus extraction using VANA was superior for recovery of sequence from DNA viruses, but its performance deteriorated as community complexity increased (Schönegger et al., 2023b).
This same study also found that ribosome-depleted RNASeq performed better than extracting using VANA or dsRNA extraction methods on datasets combined *in silico* from single-isolate studies with longer contigs, less biased results and identification of 90% of the viruses in the sample even in complex datasets, though this comparison was not made on RNA pools with many viruses physically present in the sample. A later survey found that identification was no different between physically bulked and single-plant (higher depth) sequencing, though more of the genome was recovered when RNA extracted from a single plant was sequenced (Fowkes et al., 2021b).
In any of these cases, the end result is enriching nucleotides of a type likely to contain viral sequence compared to non-viral sequence. During this thesis, these methods were not compared, but both RNA and DNA were recovered, including several complete genomes.
As well as enriching for target sequences, contamination is a concern in HTS, especially with older Illumina based methods which suffered from more frequent index-hopping, but the contamination can occur at any point in the pipeline (Kutnjak et al., 2021). Contamination due to index-hopping is of particular concern in some older SRA datasets which predate the improvements in the technology. In addition, mining of SRA data may mean that reads from associated healthy or alien controls used in the same run may not be available. This complicates the process of screening for contaminants and may mean it is not possible to cross-check at all. For example, where a virus was detected within the pooled RNA extract from plants which later tested positive individually via PCR, even at relatively low incidence (3/27 plants), contigs were obtained which represented the majority of the genome (chapter 3). A fragmentary contig was found with homology to CMV, and subsequently ruled out as present in any of the individual plants extracted via both PCR and ELISA. The CMV detection was deemed to be the result of contamination based on this, and based on comparison with reads from other samples multiplexed together on the same Illumina run. This highlights the importance of confirmatory tests using, if not an entirely different biological method (e.g. ELISA, which relies on antisera but may miss variants with particularly different coat proteins) at least a new extract. The healthy

control from the same Illumina run did not test positive for CMV, though there were CMV-positive samples from other projects on the same run which were the likely source of the CMV sequence present in the oca datasets.
The use of intentional alien controls to determine a baseline of contamination could have assisted with this conclusion (Rong et al., 2023), though the drawback would be that an additional index slot on each run cannot be used for sequencing samples, reducing the cost-effectiveness. Negative controls are already frequently excluded due to cost (Kutnjak et al., 2021).
ELISA tests for Arabis mosaic virus using commercially available antisera did not produce any positive results in extracts from the plants whose extracted RNA was pooled for HTS sequencing. Only by designing primers to the exact isolate present in the plants and then testing new RNA extracts via qRT-PCR was ArMV confirmed via a secondary method. Similarly, the yacón plants which tested positive for PYV via HTS were not detected via qRT-PCR on either the extract subjected to HTS or a secondary extract. Conventional PCR of the original extracts confirmed the infection, but did not produce a positive result in the second extract one of the two plants. This indicates some of the difficulty in using confirmatory, targeted test to confirm infection of a plant with a divergent isolate first detected in a plant via a non-targeted method.

7.10. Conclusion

Some viruses are difficult to diagnose without molecular methods, which are expensive, and may increase the price of plant material that a grower could select from when choosing to grow a crop. This increased price may be prohibitive or at least unfavourable for the hobbyist grower. Cheap, untested seed thus circulates to serve an informal economy selling what is, in Europe, an unusual crop. The present work that many viruses are associated with ARTCs intended for planting which are illegally imported into the UK. Some of these viruses were first detected and described in the present work, while others were already known to science. In addition, some of these putative novel viruses, as with many other viruses first discovered via HTS, have properties which make it challenging to produce data on their biology using classical virological methods. Plant health management requires 'knowing the enemy'. In the case of the viruses detected infecting oca, these 'enemies' have been assigned their roles based on their probable taxon: that they are viruses within known pathogenic genera. Experiments on yield reduction in mashua (Guimaraes and Flores, 2005) and ulluco (Lizárraga et al., 2001) and the detection of potentially destructive viruses such as ArMV, TuMV and PYV in ARTC samples do suggest that ARTCs pose plant health challenges, both to the both the crops themselves and other crops presently grown in Europe more widely.
A pest risk assessment as per the ISPMs can form a framework for investigating a virus but, as has been shown, this process may not be suited to all pathogens. Clive Brasier in a 2005 article cited over-reliance on pest-specific PRAs as one of the challenges threatening biosecurity, especially as harmful organisms may be overlooked by NPPOs is they are absent from quarantine lists (Brasier, 2005).

A ’virus’ might be considered as a taxonomic unit, as a cycle or process, more than a stable object (Dupré and Guttinger, 2016). High throughput sequencing and the associated developments of data-driven discovery tie intimately into these fundamental topics (Lauber and Seitz, 2022). Increasingly the notion of ’a virus sequence’ is being determined, without biological properties that derive from the translation of that sequence into proteins within a living host cell, while our ability to infer phenotype from virus genes is still in its infancy (Simmonds et al., 2017). This can complicate the distinction made between ’sequence’ and ’virus’ in some cases, as often it is sequence being reported, not a virus, with phylogeny often becoming a tree of genes rather than of viruses as a whole (Gibbs, 2020). Thus the ’biological desert’ in novel virus findings remains somewhat intractable, as demonstrated here by the difficulty in obtaining biological information such as an experimental host range (Hou et al., 2020). It is the virus, the biological entity, which infects the cell, which the sequence is only an indicator of, and this virus as an actual object with properties is the basis for the risk assessment. The difficulty is influenced here by different factors:

1. the data level, relating to information sources and contamination.
2. the virus, both the difficulty of assigning taxonomy to this variable group and our emerging understanding of their interactions with hosts.
3. the practical, an issue of time, space and money.

The crucial matter of ’causation’ hovers between the three. While nothing can be done to change how viruses operate as biological entities, resources can (and must) be allocated wisely. Data can be interpreted, shared and re-used with care. Once acknowledged, data gaps can be filled with judicious studies, such as surveys, allowing some viruses to be lowered in priority to avoid wasting resources (Fontdevila Pareta et al., 2023).
It remains to be seen whether these plants, in ways so similar to potato, will follow an arc of progress like their well-studied cousin. Will the existence of high speed internet trade accelerate them down a similar path, or will that rapid, long-distance movement alter the evolutionary history of their virome in entirely different ways?

Appendix A. R Scripts Provided by Roy Macarthur

Binomial Distribution Test for Host Range Study

```
library(VGAM)

previous.x<-1
previous.n<-1000
alpha<-previous.x+1
beta<-previous.n-previous.x+1
number.of.new.tests=c(100:110,150,200,250,400)

prob.all.negative=pbetabinom.ab(q=0,size=number.of.new.tests,shape1=alpha,shape2=beta)

prob.all.negative.pc=round(100*prob.all.negative,1)
data.frame(number.of.new.tests,prob.all.negative.pc)
```

A.1. Effect of Nepovirus Subgroup on Host Range

This script was used to generate Figure 5.2 on Page 124.

```
library(reshape2)
## Warning: package 'reshape2' was built under R version 4.1.3
library(ggplot2)
## Warning: package 'ggplot2' was built under R version 4.1.3
library(plyr)
## Warning: package 'plyr' was built under R version 4.1.3
library(lme4)
## Warning: package 'lme4' was built under R version 4.1.3
## Loading required package: Matrix
## Warning: package 'Matrix' was built under R version 4.1.3
library(effects)
## Warning: package 'effects' was built under R version 4.1.3
## Loading required package: carData
## Warning: package 'carData' was built under R version 4.1.3
```

```
## lattice theme set by effectsTheme()
## See ?effectsTheme for details.
data<-read.csv("observations.csv",na.strings = "")
the.groups<-read.csv("groups.csv")
data<-melt(data,id.vars=c("Genus","Species","Common.Name"),value.name =
    "observation",variable.name = "virus")
data<-merge(data,the.groups)

data<-ddply(data,c("Subgroup","Genus","Species"),
            observations=sum(!is.na(observation)),
            positive=sum(observation=="Y",na.rm=TRUE),
            summarise)
data<-data[-which(data$observations==0),]
model<-glmer(cbind(positive,observations-positive)~Subgroup+(1|Species),family=binomial(link
    = "logit"),data=data)
print(summary(model))
```

Appendix B. Table of the Host Range of Nepoviruses

ArMV	Arabis mosaic virus	**PVB**	Potato virus B
AVA	Arracacha virus A	**TBRV**	Tomato black ring virus
AARSV	Artichoke Aegean ringspot virus	**AYRSV**	Artichoke yellow ringspot virus
CsAIV	Cassava American latent virus	**ALRSV**	Apricot latent ringspot virus
GDefV	Grapevine deformation virus	**BRV**	Blackcurrant reversion virus
GFLV	Grapevine fanleaf virus	**CGMV**	Cassava green mottle virus
OLRSV	Olive latent ringspot virus	**BLMoV**	Blueberry leaf mottle virus
PBRSV	Potato black ringspot virus	**BLSV**	Blueberry latent spherical virus
RpRSV	Raspberry ringspot virus	**CLRV**	Cherry leaf roll virus
MMMoV	Melon mild mottle virus	**CYMV**	Chicory yellow mottle virus
AeRSV	Aeonium ringspot virus	**GBLV**	Grapevine Bulgarian latent virus
MMLRaV	Mulberry mosaic leaf roll-associated virus	**GTRSV**	Grapevine Tunisian ringspot virus
TRSV	Tobacco ringspot virus	**HLRSV**	Hibiscus latent ringspot virus
AILV	Artichoke Italian latent virus	**LALV**	Lucerne Australian latent virus
BRSV	Beet ringspot virus	**MLRV**	Myrobalan latent ringspot virus
CNV	Cacao necrosis virus	**CawYV**	Carraway yellows virus
CCLV	Crimson clover latent virus	**PRMV**	Peach rosette mosaic virus
CNSV	Cycas necrotic stunt virus	**PVU**	Potato virus U
GARSV	Grapevine Anatolian ringspot virus	**ToRSV**	Tomato ringspot virus
CGMV	Grapevine chrome mosaic virus	**SLSV**	Soybean spherical latent virus
MRSV	Mulberry ringspot virus		

Key:

Y	Host	**t**	Nematode
N	Non-host	**o**	Nematode only
B	Sources disagree	**x**	Asymptomatic
A	Natural host	**s**	Seed-transmitted
		l	Local lesion
		p	Purification
		m	Maintenance
		c	Confirmed non-host
		e	Mechanical only

Family	Species	ArMV	AVA	AARSV	CsAlV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Acanthaceae	*Thunbergia alata*													Nc																											N	
Adoxaceae	*Sambucus canadensis*																														Y									A	
Adoxaceae	*Sambucus nigra*	Y								Y														A							A										
Adoxaceae	*Sambucus racemosa*																														Y										
Adoxaceae	*Sambucus spp.*													Nt																	A										
Aizoaceae	*Tetragonia expansa*		Y		N		N							Y					Y					Y			Nc												Y	Y	
Aizoaceae	*Tetragonia tetragonioides*	Y	Y						Y					Y	Y		Y	N	Y			N		Y				Y									Y				N	
Alstroemeriaceae	*Alstroemeria sp.*	A																																							
Amaranthaceae	*Amaranthus caudatus*	B	Y				Y		Y					Y	Y									Y							N									Y	
Amaranthaceae	*Amaranthus cruentus*													Y																											
Amaranthaceae	*Amaranthus hybridus*													Y	Y																	Y									
Amaranthaceae	*Amaranthus palmeri*													Y																											
Amaranthaceae	*Amaranthus paniculatus*						Y							Y																											
Amaranthaceae	*Amaranthus patulus*						Y																																		
Amaranthaceae	*Amaranthus retroflexus*	N					Y							Y	Y							N		Y									Y							Y	
Amaranthaceae	*Amaranthus spinosus*																																							A	
Amaranthaceae	*Amaranthus tricolor*						Y																		Y															Nc	
Amaranthaceae	*Amaranthus viridis*														Y																										
Amaranthaceae	*Atriplex hortensis*	Y												Y	Y									Y							Yp		Y								
Amaranthaceae	*Beta macrocarpa*																														Y										
Amaranthaceae	*Beta vulgaris*	Bt	Y				Y		Y	Y				Bl		Yt	Y	N	Y					B	Y						Y	Y	Y							Y	Y	
Amaranthaceae	*Celosia argentea*	B												Y	Y									Y	Y						Y	Y								N	
Amaranthaceae	*Celosia cristata*						N								Y				Y												Y	Y	Y								
Amaranthaceae	*Celosia plumosa*														Y																										
Amaranthaceae	*Celosia spp.*														Y																										
Amaranthaceae	*Chenopdoium opulifolium*														Y																										
Amaranthaceae	*Chenopodium album*	Y			Y		Ym		Yl					Y				Y				Y		B					Y	Nc	Y	Y			Y				Y	Y	
Amaranthaceae	*Chenopodium amaranticolor*	B	Y		Y	Y	Yms	Y	Yl	Yl	Y			Yl		Yt	Y	Y	Bs	Y	Y	Y		Yl	Y		Yl	Y	Y		Ys	Yl	Y	Yl	Ys	Y	Y		Ym	Y	Yl	
Amaranthaceae	*Chenopodium ambrosioides*	Y					Y																		Y								Y								
Amaranthaceae	*Chenopodium bonus*						N																																		
Amaranthaceae	*Chenopodium bushianum*					Y														Y																					
Amaranthaceae	*Chenopodium capitatum*						Y		Y						Y															Y			Y							Yx	Y	
Amaranthaceae	*Chenopodium ficifolium*														Y									Y																Y		
Amaranthaceae	*Chenopodium foetidum*	Y				Yx	Y		Y						Y			N		Y	Y			Y		Y			Y	Y						Y				Y	Y	
Amaranthaceae	*Chenopodium foliosum*		Y				Y		Y						Y																										
Amaranthaceae	*Chenopodium giganteum*						Y			Y				Y	Y																										Y	
Amaranthaceae	*Chenopodium glaucum*																																						Y		
Amaranthaceae	*Chenopodium hybridum*	Yt					Y								Y																		Y							Y	
Amaranthaceae	*Chenopodium murale*	Yt	Y				Y		Y			Yl			Y			N	Yl			Y		Y	Y	Y	Yl	Y	Y		Y		Y		Yl	Yl				Y	Y	
Amaranthaceae	*Chenopodium polyspermum*						Y																																		
Amaranthaceae	*Chenopodium quinoa*	Bt	Ym			Y	Yms	Ym	Yl	Y	Yl	Ypl		Y	Ym		Y	Y	B	Y	Yl	Y		Yb	Ypl	Y	Ym	Yl	Ym	Yp	Yp	Y	Ypl	Y	Ys	Ym	Ypl		Ym	Y	Yl	
Amaranthaceae	*Chenopodium serotinum*																		Ys																						
Amaranthaceae	*Chenopodium urbicum*														Y																										
Amaranthaceae	*Chenopodium viride*														Y																										
Amaranthaceae	*Chenopodium vulvaria*														Y																										
Amaranthaceae	*Cichorium endiva*													Y											Y																
Amaranthaceae	*Cichorium intybus*													Y	Y										Nc								A									

Family	Species	AfMV	AVA	AARSV	CsAlV	GDeIV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Amaranthaceae	*Gomphrena celosioides*																																									A
Amaranthaceae	*Gomprena globosa*	Bc	Y		N	Yx	Ym	Yl	Y	Y	Y			Yln	Y		Y	N	Y			N		Y	Ym	N	Y	Y		B	Y		Y	Yx	Yl	Ym			Yl	Y	Y
Amaranthaceae	*Spinacia oleracea*	Y								Yb	Y			Y	Y	Yt	N		B			N		Y	Y		Yx									Y				Yx	Nc
Amaranthaceae	*Taraxacum officinale*	Y												A										A																As		Yt
Amaryllidaceae	*Allium ampeloprasum*																							A																		
Amaryllidaceae	*Allium cepa*													Nc										A																		Nc
Amaryllidaceae	*Allium porrum*																							A																		
Amaryllidaceae	*Allium. spp.*																							A																		
Amaryllidaceae	*Narcissus pseudonarcissus*	Y								A				A										A																		A
Amaryllidaceae	*Narcissus spp.*	Y						Y						Y																												
Apiaceae	*Aegopodium podagraria*																														Y											
Apiaceae	*Ammi majus*														Y									Y							Y	Y										
Apiaceae	*Apium graveolens*	Yx							Y					Y							A			A				Y			Y										Y	
Apiaceae	*Arracacia xanthorrhiza*		Y						A																																	
Apiaceae	*Arethum graveolens*																								Y																	
Apiaceae	*Arthriscus cerefolium*		Y											Y																											Y	
Apiaceae	*Carum carvi*																																A									
Apiaceae	*Carum segetum*	Y																																								
Apiaceae	*Cnidium officinale*																		A																							
Apiaceae	*Coriandrum sativum*	Yx	Y																					Y																		
Apiaceae	*Daucus carota*	Bc	N						Y					Nc			N					N		A				Y		Nx											Nc	
Apiaceae	*Eryngium sp.*	A																																							A	
Apiaceae	*Eryngium creticum*																								Y																	
Apiaceae	*Foeniculum vulgare*																																									
Apiaceae	*Heracleum sphondylium*															A																										
Apiaceae	*Pastinaca sativa*													Nx			N							A						Nx												
Apiaceae	*Pastinaca anisum*																														Nx											
Apiaceae	*Petroselinum crispum*		N																					A							A										
Apiaceae	*Petrosilenum sativum*																														Y											
Apocyanaceae	*Apocyanum cannabinum*													A											Yl																	
Apocyanaceae	*Araujia sericifera*																																									
Apocyanaceae	*Catharanthus roseus / Vinca rosea*								Y					B	Nx					Nc			N					Y	N		Y											B	
Apocyanaceae	*Vinca minor*	A																																								
Araceae	*Arum sp.*	A																																								
Araceae	*Spathiphylum sp.*																																								A	
Arialaceae	*Hedera helix*	A																																								
Arialaceae	*Schefflera arboricola*																																								A	
Arialaceae	*Schefflera sp.*													Y																												
Arialaceae	*Trachymeme coerulea*													Y																												Nc
Aristolochiaceae	*Aristolochia clematitis*						A																																			
Asparagaceae	*Asparagus officindalis*	Y												Nc										A																		Nc
Asparagaceae	*Camassia cusickii*	A																																								
Asparagaceae	*Camassia quamash*	A																																								
Asparagaceae	*Hosta fortunei*	A																																								
Asparagaceae	*Hosta spp.*	Y												Y										A																		A
Asparagaceae	*Muscari sp.*	A																																								
Asphodelaceae	*Hemerocallis sp.*													Y																												

Family	*Species*	ArMV	AVA	AARSV	CsAlV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Asphodelaceae	*Kniphofia uvaria*													Nc																												Nc	
Asteraceae	*Acanthospermum hispidum*																																									A	
Asteraceae	*Achillea millefolium*													Y										Y																			
Asteraceae	*Ageratum conyzoides*													Y										N																		Nc	
Asteraceae	*Ageratum houstonianum*													Y																													
Asteraceae	*Amberboa moschata*													Y																													
Asteraceae	*Ambrosia artemisifolia*													Y																													
Asteraceae	*Ambrosia trifida*													Y																													
Asteraceae	*Arctium lappa*													A		A																											
Asteraceae	*Argyranthemum frutescens*													Y																													
Asteraceae	*Aster laevis*													Y																													
Asteraceae	*Aster spp.*																							N								Y										Nc	
Asteraceae	*Bellis perennis*	Y												Y										Y																			
Asteraceae	*Bidens discoidea*													Y																													
Asteraceae	*Bidens pilosa*																																									A	
Asteraceae	*Calendula officinalis*	Y												B	Nx									Y	Y								Y										
Asteraceae	*Callistephus chinensis*													Y											Y																		Nc	
Asteraceae	*Carduus pychnocephalus*																								Y																		
Asteraceae	*Carthamus tinctorius*																								Y																		
Asteraceae	*Chrysanthemum × morifolium*													Y										N																		A	
Asteraceae	*Chrysanthemum carinatum*													Nx											Y																		
Asteraceae	*Chrysanthemum indicum*																																									A	
Asteraceae	*Chrysanthemum segetum*																								Y																		
Asteraceae	*Chrysanthemum sp.*													Y																													
Asteraceae	*Cirsium arvense*													Nx											Y																		
Asteraceae	*Cirsium vulgare*													A																													
Asteraceae	*Cosmos diversifolius*													Nx																													
Asteraceae	*Crepis neglecta*														A																												
Asteraceae	*Crepis tectorum*													Y																													
Asteraceae	*Cynara cardunculus / Cytra scolymus*			Y						A					Yt									Y	Y																		
Asteraceae	*Dahlia merckii*													Nx																													
Asteraceae	*Dahlia pinnata*																							N																			
Asteraceae	*Dahlia spp.*	A																						N																		A	
Asteraceae	*Echinacea purpurea*													Y																													
Asteraceae	*Emilia coccinea*														Y																													
Asteraceae	*Emilia flammea*																															Yl											
Asteraceae	*Emilia sagittata*	Y																						Y																			
Asteraceae	*Erigeron annuus*													A																													
Asteraceae	*Erigeron canadensis*													Y											Y																		
Asteraceae	*Eupatorium capillifolium*													A																													
Asteraceae	*Eupatorium purpureum*													Y																													
Asteraceae	*Galinsoga parviflora*													Y																													
Asteraceae	*Galinsoga parviflora*																						N																			
Asteraceae	*Galinsoga quadriradiata*													Y																													
Asteraceae	*Gebera spp.*																			Nc																							
Asteraceae	*Gerbera jamesonii*																							A																			
Asteraceae	*Gerbera sp.*													Y																													

Family	Species	ArMV	AVA	AARSV	CsAIV	GDeFV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Asteraceae	*Helenium amarum*													A	Y																											
Asteraceae	*Helianthus annuus*													Nc	Nx				Nc					Y	Y						Yl											
Asteraceae	*Helminthia echioides*														A										Nc																	
Asteraceae	*Hypochaeris achyrophorus*														Y																											
Asteraceae	*Hypochaeris aetnensis*														A																											
Asteraceae	*Lactuca sativa*	B							Y					Y			N		Nc					Yt	Y			Y					Y								B	
Asteraceae	*Lactuca serriola / Lactuca scariola*													Y														Y														A	
Asteraceae	*Lactuca virosa*														A																											
Asteraceae	*Laggeria aurita*																							Yx																	A	
Asteraceae	*Lapsana communis*																																									
Asteraceae	*Leucanthemum vulgare*																							Y																	A	
Asteraceae	*Liatris spicata*																							Nc																		
Asteraceae	*Liatrus sp.*	A																						A																	A	
Asteraceae	*Scolymus maculatus*																								Y																	
Asteraceae	*Senecio vulgaris*	Y												Y										Y					Y													
Asteraceae	*Solidago x hybrida*																							Ym																	
Asteraceae	*Sonchus arvensis*													Y	A									Nc																	
Asteraceae	*Sonchus asper*																							A																	
Asteraceae	*Sonchus oleraceus*	Nc					A							Y										Y	Y																Nc	
Asteraceae	*Sonchus spp.*														A																											
Asteraceae	*Symphyotrichum pilosum*													A										Bc	Y																A	
Asteraceae	*Tagetes erecta*													Y										Y	Y																	
Asteraceae	*Tagetes patula*																																								Nc	
Asteraceae	*Tagetes signata*															A																										
Asteraceae	*Tanacetum vulgare*													Y															N	N												
Asteraceae	*Tithonia rotundifolia / Tithonia speciosa*																								Y				N												
Asteraceae	*Torilis nodosa*													Nx																											
Asteraceae	*Tragopogon porrifolius*	A															A																								
Asteraceae	*Tussilago farfara*														A										Y																
Asteraceae	*Urospermum dalechampii*													Yln																											
Asteraceae	*Verbesina alternifolia*													A																											
Asteraceae	*Xanthium strumarium*	Y	N											Yln	Nx					B			N		Y	Y			N			Y							Y		Y	
Asteraceae	*Zinnia elegans*													A																											
Annonaceae	*Asimina triloba*													Y																										Y	
Balsaminaceae	*Impatiens balsamina*																																							A	
Balsaminaceae	*Impatiens sp.*													A																											
Balsaminaceae	*Impatiens walleriana*													Y																										Y	
Begoniaceae	*Begonia cuculatta*															A																									
Begoniaceae	*Begonia rieinifolia*																																								
Begoniaceae	*Begonia sp.*																							Yo																	
Berberidaceae	*Berberis vulgaris*													Nx																											
Betulaceae	*Betula nigra*																																								
Betulaceae	*Betula pendula*	A																																							
Betulaceae	*Betula pubescens*																																								
Betulaceae	*Betula spp.*																																								
Betulaceae	*Betula verrucosa*																																								

Family	Species	AfMV	AVA	AARSV	C_{SA}IV	GDeFV	GFLV	OLRSV	PBRSV	R_{PR}SV	MMMoV	AeRSV	MMLR$_{aa}$V	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	C_{aw}YV	PRMV	PVU	TaRSV	SLSV
Betulaceae	*Carpinus betulus*													Nc																Y												
Bignoniaceae	*Bignonia capreolata*													Y																Y												Nc
Bignoniaceae	*Incarvillea delavayi*													Nc																												Nc
Boraginaceae	*Archusa azurea*													Y																												Y
Boraginaceae	*Cynoglossum amabile*													Y																												Nc
Boraginaceae	*Heliotropum arohorescens*													Y																												Y
Boraginaceae	*Myosotis arvensis*																								A																	
Boraginaceae	*Myosotis sylvatica*	Y												Nc											Y																	Nc
Boraginaceae	*Nemophila insignis*													Y																												Y
Boraginaceae	*Phacelia campanularia*													Y																												Nc
Boraginaceae	*Pulmonaria officinalis*																													N											
Brassicaceae	*Amoracia rusticana*	Y												A																											
Brassicaceae	*Arabidopsis thaliana*																														Ys										
Brassicaceae	*Arabis alpina*																																								Nc
Brassicaceae	*Arabis hirsuta*	Y																							Y																
Brassicaceae	*Barbarea sp.*													Y																											
Brassicaceae	*Betertetoa incana*																														A										
Brassicaceae	*Brassica alba*													Nx				N	N	Nc										N											
Brassicaceae	*Brassica campestris*	Y	Y											N											PVO																	
Brassicaceae	*Brassica juncea*	Y																																							
Brassicaceae	*Brassica napus*													Nx											Y																Nc
Brassicaceae	*Brassica nigra / Sinapis nigra*																																								
Brassicaceae	*Brassica oleracea*	Y	Y											Nx						Nc					A					N											N
Brassicaceae	*Brassica rapa*													Bc											Nc	Nc				Y											Bl
Brassicaceae	*Brassica rapa subsp. rapa*													Y		Yt							N		Yb																
Brassicaceae	*Brassica rapa var. pekinensis*																																							N	
Brassicaceae	*Brassica rapa var. pervitidis*																														Y										
Brassicaceae	*Brassica spp.*																								A																
Brassicaceae	*Capsella bursa-pastoris*	Y								A				Y											Yb																
Brassicaceae	*Cheiranthus cheiri*													Y											Yx					Y										
Brassicaceae	*Eruca sativa / Rucola sativa*														Y															Nx										
Brassicaceae	*Erysimum cheiranthioides*																								Nc																
Brassicaceae	*Iberis saxatilis*									Y																															
Brassicaceae	*Lepidium densiflorum*													A																											A
Brassicaceae	*Lepidium didymum*													A																											A
Brassicaceae	*Lepidium draba*																								Y																
Brassicaceae	*Lobularia maritima*													Y																											Y
Brassicaceae	*Malconia maritima*													Nx																											
Brassicaceae	*Matthiola incana*	N																							Y																Nc
Brassicaceae	*Matthiola spp.*																								N																
Brassicaceae	*Raphanus sativus*													Nx			N			Nc						Nc					Nc	Y									B
Brassicaceae	*Rhamphospermum arvense*													A																											A
Brassicaceae	*Sinapis alba*																								Y																
Brassicaceae	*Sinapis sp.*																									Y															
Brassicaceae	*Sisymbrium officinale*																									Y															
Buxaceae	*Buxus sempervirens*	Y																																							
Campanulaceae	*Campanula medium*													Y																											Y
Campanulaceae	*Campanula spp.*																								Y																

Family	Species	ArMV	AVA	AARSV	C_sAIV	GD_eFV	GFLV	OLRSV	PBRSV	R_pRSV	$MMLO_aV$	Ae_RSV	$MMLR_aV$	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	$BLMO_aV$	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	L.ALV	MLRSV	$C_{aw}YV$	PRMV	PVU	ToRSV	SLSV
Campanulaceae	*Lobelia cardinalis*													Nc																											Nc	
Campanulaceae	*Lobelia erinus*	Y																						Y							Y											
Campanulaceae	*Lobelia spp.*	Y																																								
Campanulaceae	*Lobelia tenuior*								Y					Y																											Nc	
Campanulaceae	*Platycodon grandiflorus*								Y					Y																											Nc	
Cannabinaceae	*Humulus lupulus*	Y																																							
Caprifoliaceae	*Centranthus ruber*													Nx																											
Caprifoliaceae	*Dipsacus sylvestris*																							Yx																	
Caprifoliaceae	*Scabiosa atropurpurea*								Y					Y																											Nc	
Caprifoliaceae	*Scabiosa japonica*													Nc																											Nc	
Caprifoliaceae	*Valerinaa officinalis*													Nc																	Nx										Nc	
Caprifoliaceae	*Wiger sp.*									Y																																
Caricaceae	*Carica papaya*													A																												
Caryophyllaceae	*Cerastium vulgatum*									A														A							N	Nx									
Caryophyllaceae	*Dianthus barbatus*	Y							Y					Nx										N	Y																Nc	
Caryophyllaceae	*Dianthus caryophyllus*	Y												Y											Y							Nx									Y	
Caryophyllaceae	*Dianthus chinensis*																			Nc																						
Caryophyllaceae	*Dianthus spp.*																																									
Caryophyllaceae	*Gypsophila elegans*								Y					Y																											Y	
Caryophyllaceae	*Gypsophila paniculata*																																							Ap	
Caryophyllaceae	*Gypsophila perfoliata*														Nx																										
Caryophyllaceae	*Lychnis coronaria*													Y																											Nc	
Caryophyllaceae	*Saponaria vaccaria*																								Y			N													
Caryophyllaceae	*Silene chalcedonica*	A												Y												Y															Y	
Caryophyllaceae	*Silene sp.*																							Yo	Y																	
Caryophyllaceae	*Spergula arvensis*									A																																
Caryophyllaceae	*Stellaria media*	Y							Y	As				A											Y						Y										
Celastraceae	*Celastrus orbiculatus*															A								A																
Celastraceae	*Euonymus alatus*	Y																													Y										
Celastraceae	*Euonymus europaeus*																														Y										
Celastraceae	*Euonymus fortunei*																														Y										
Celastraceae	*Euonymus kiautschovicus*																														Y										
Cleomeaceae	*Cleome spinoa*													Y																											Y	
Colchicaceae	*Colchicum sp.*													Y																											Y	
Colleminaceae	*Commellina erecta*																																							A	
Colleminaceae	*Commellina spp.*																														Y										
Colleminaceae	*Commellina coelestis*																														N										
Colleminaceae	*Tinantia erecta*																														Y										
Colleminaceae	*Tradescantia fluminensis*													Nx																											
Colleminaceae	*Tradescantia zebrina*													Nx																											
Convulvulaceae	*Convolvulus tricolor*													Y																											Y	
Convulvulaceae	*Ipomoea batatas*													A																											A	
Convulvulaceae	*Ipomoea eriocarpa*																																								A	
Convulvulaceae	*Ipomoea nil*													Y																												
Convulvulaceae	*Ipomoea purpurea*													Y																												
Convulvulaceae	*Ipomoea setosa*																																			N						
Cornaceae	*Cornus florida*																															A										
Cornaceae	*Cornus racemosa*													A																												

Family	Species	ArMV	AVA	AARSV	CsAIV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Cornaceae	*Cornus sericea*													A																											
Cornaceae	*Cornus spp.*													A																										A	
Crassulaceae	*Aeonium spp.*											Y		A																											
Crassulaceae	*Kalanchoe pinnata / Bryphyllum pinnatum*													Nx																											
Cupressaceae	*Chamaecyparis lawsoniana*	Y																						A																	
Curcurbitaceae	*Benincasa hispida*										Y																														
Curcurbitaceae	*Citrullus lanatus*										N	Y		Ao											Y					PVO											
Curcurbitaceae	*Citrullus vulgaris*													Nb																											Nc	
Curcurbitaceae	*Cucumis melo*	Y									Y	Y		Yp					Y					Y	Y					Y	Y									B	
Curcurbitaceae	*Cucumis sativa*	Y	Y				Y	Y	Y	Y	N	Y		Y	Y		Y	N	Nc		Y	N		Yb	Ylp	Nx	Y	Y	Y		Y	Y	Y	Y		Y			Yo	Y	b	
Curcurbitaceae	*Cucurbita moschata*													Y					Nc																						
Curcurbitaceae	*Curcurbita maxima*	Y					Y							Y											Nc				N		Y					N				Y	
Curcurbitaceae	*Curcurbita pepo*	Y	Y								N	Y		Y	Ym				Y			N		Y	Y				N		B	Yp	Y	Y						Y	
Curcurbitaceae	*Curcurbita spp.*									Y																															
Curcurbitaceae	*Lagenaria leucnatha*													Y																											
Curcurbitaceae	*Lagenaria siceraria / Lagenaria leucantha*						A				N			Y																											
Curcurbitaceae	*Luffa acutangula*																							Y																	
Curcurbitaceae	*Luffa cylindrica*													Y																											
Curcurbitaceae	*Momordica balsamina*																												N		N									
Curcurbitaceae	*Sedum acre*													Nc																											Y	
Cycadaceae	*Cycas revoluta*																		Y																						
Equisetaceae	*Equisetum arvense*													Y																											
Ericaceae	*Vaccinium augustifolium*																												A											
Ericaceae	*Vaccinium corymbosum*													At															Y	A									A		At	
Ericaceae	*Vaccinium darrowii*																														A									
Ericaceae	*Vaccinium myrtilloides*																												A											A	
Ericaceae	*Vaccinium spp.*													Y																											A	
Euphorbiaceae	*Euphorbia cyparissias*																							Nc																	
Euphorbiaceae	*Euphorbia prunifolia*				N																																			
Euphorbiaceae	*Euphorbia supina*																																							A	
Euphorbiaceae	*Manihot esculenta*				Yp																							A												
Euphorbiaceae	*Mercurialis perennis*	A																							Y																
Euphorbiaceae	*Ricinus communis*													Y														Y												Y	
Fabaceae	*Arachis hypogaea*													Nc	Nx				Y									N			N					N				Nc	
Fabaceae	*Cajnatus cajnat*																																			Y					
Fabaceae	*Canavalia ensiformis*																								Yl																
Fabaceae	*Cicer arietinum*																								Yl											Y					
Fabaceae	*Clitoria terantea*		Y																																						
Fabaceae	*Crotalaria juncea*								Y																																
Fabaceae	*Crotalaria spectabilis*													Y																	Y					N					
Fabaceae	*Cyamopsis tetragonoloba*	Y	Y						Y					Yt															Y		Y										
Fabaceae	*Dolichos biflorus*								Y						Y																											
Fabaceae	*Dolichos sp.*																								Y																
Fabaceae	*Glycine max*	Y					Y							Ysn			Y	N	B			Y		Y	Y						A					N				B	A
Fabaceae	*Hedysarum coronarium*																															Nx									

Family	Species	ArMV	AVA	AARSV	CsAIV	GDeFV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Fabaceae	*Lablab purpureus / Dolichos lablab*								Y					Y	Y										Y											N						
Fabaceae	*Laburnum × watereri*	A																						N	Y		Nc	Y			Y											
Fabaceae	*Lathyrus odoratus*	Yx												Y					B					N	Y																		
Fabaceae	*Lens culinaris*													Y										Y																			
Fabaceae	*Lens esculenta*																								Y																	
Fabaceae	*Lotus corniculatus*													Y																							N						
Fabaceae	*Lupinus albus*													Y											Y							Y					Y						
Fabaceae	*Lupinus angustifolia*																																				Y						
Fabaceae	*Lupinus hirsutus*	Y												Y																													
Fabaceae	*Lupinus luteus*																							Nc																			
Fabaceae	*Lupinus polyphyllus*													Y																			Nx										
Fabaceae	*Lupinus spp.*													A						Nc																							
Fabaceae	*Macroptilium lathyroides*													Y																													
Fabaceae	*Medicago sativa*													Y					N	A					A							Ys										
Fabaceae	*Medicago spp.*																															Y										
Fabaceae	*Melilotus albus*													Y										N																		
Fabaceae	*Melilotus officinalis*	Y												Y												Y																	
Fabaceae	*Melilotus sp.*																																									
Fabaceae	*Mucuna sp.*													Y																													
Fabaceae	*Phaseolus acutifolius*		N																																							
Fabaceae	*Phaseolus coccineus / Phaseolus multiflorus*	Y													Y																												
Fabaceae	*Phaseolus lunatus*													Y	Nx									A																			
Fabaceae	*Phaseolus spp.*																																									
Fabaceae	*Phaseolus vulgaris*	Y	N					Y		Y	N			B	Yml	Y	Ym	N	Bl		Yp	Y		Y			Nc	Y	Y		Y	Y					Y			N	B	
Fabaceae	*Pisum sativum*	Y							Y					Y	Yml		N	N	Bl			Y		Y	Y			N			Ys	Nx					Ym	Ym		Yc	Yc
Fabaceae	*Pueraria montana*													A										A																	
Fabaceae	*Robinia pseudoacacia*															Y																										
Fabaceae	*Senna obtusifolia*																																								A	
Fabaceae	*Senna occidentalis /Cassia occidentalis*													Yt															N												Nc	
Fabaceae	*Senna tora / Cassia tora*																															N									Nc	
Fabaceae	*Sesbania exaltata*																															Y										
Fabaceae	*Sophora microphylla*													A															N													
Fabaceae	*Tifolium hybridum*													N																													
Fabaceae	*Trifolium pratense*													Nc					N	Nc				N													Y					A	
Fabaceae	*Trifolium spp.*																																				Y					A	
Fabaceae	*Trifolium repens*	At												Nc			N	N	Nc					A													Y					A	
Fabaceae	*Trifolium subterraneum*																	N																			Y						
Fabaceae	*Trifolium incarnatum*													Y				A																			Y	Y					
Fabaceae	*Vicia faba*	Y						Y			N			B			N	N	Nc				N	B	Y			Y		Nx							Y				Y	
Fabaceae	*Vicia sativa*													Y											Y													Y					
Fabaceae	*Vicia villosa*													N																													
Fabaceae	*Vigna angularis*													Y																													
Fabaceae	*Vigna cylindrica*								Y						Y																										Yx		
Fabaceae	*Vigna radiata / Phaseolus aureus*		N						Y					Y											Y							Y	Nx	N								

Family	Species	ArMV	AVA	AARSV	CsAlV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Fabaceae	*Vigna sesquipidalis*																		B																							
Fabaceae	*Vigna umbellata*													Y																												
Fabaceae	*Vigna unguiculata / Vigna sinensis*	Y	N				Y		Y	Y	Y	Y		Bl			Y	N			Y	Yml		Y	Y	Nx			PVO		Y	Y				PVO				Yx	Yl	
Fagaceae	*Fagus sylvatica*																							A							Y										
Gentianaceae	*Gentian sp.*															A																									
Geraniaceae	*Erodium cicutarium*													Y																												
Geraniaceae	*Geranium dissectum*																							A																		
Geraniaceae	*Geranium molle*																								Y																	A	
Geraniaceae	*Pelargonium × peltato-zonale*																							Y																		
Geraniaceae	*Pelargonium × peltatum*																							Y																	A	
Geraniaceae	*Pelargonium × zonale*														Y									Y																		
Geraniaceae	*Pelargonium domesticum*													Nx																												
Geraniaceae	*Pelargonium hortorum*													A																											Y	
Geraniaceae	*Pelargonium spp.*													A	A									A							Y										A	
Geraniaceae	*Pelargonium zonale*														Y																											
Gesneriaceae	*Gloxinia spp.*																							Yx																		
Gesneriaceae	*Sinningia speciosa*													Nc																											Y	
Grossulariaceae	*Ribes alpinum*																										A															
Grossulariaceae	*Ribes aureum*																										A															
Grossulariaceae	*Ribes fasciculatum*																										A															
Grossulariaceae	*Ribes fragrans*																										A															
Grossulariaceae	*Ribes nigrum*	Y								A				A										A			Y				A										A	
Grossulariaceae	*Ribes rubrum*									A														A			A				A											
Grossulariaceae	*Ribes sachalinensis*																							A																		
Grossulariaceae	*Ribes sanguineum*																							A																		
Grossulariaceae	*Ribes spicatum*																										A															
Grossulariaceae	*Ribes spp.*	Y						Y		Y														A			Y															
Grossulariaceae	*Ribes uva-crispa*																																								A	
Hrdrangeaceae	*Hydrangea macrophylla*													Y																	Y										
Hrdrangeaceae	*Hydrangea aspera*	A																																							
Hrdrangeaceae	*Hydrangea paniculata*													Y																											
Iridaceae	*Crocus chrysanthus*	A																																							
Iridaceae	*Crocus spp.*	Y														A																										
Iridaceae	*Crocus vernus*	A																																							
Iridaceae	*Gladiolus hortulAtum*																														A										
Iridaceae	*Gladiolus spp.*	A												A	Y				A					A																	A	
Iridaceae	*Iris ensata*													Y																											
Iridaceae	*Iris germanica*													Y																											A	
Iridaceae	*Iris hollandica*	A												A																											A	
Iridaceae	*Iris sibirica*													Y																											
Iridaceae	*Iris spp.*													A																											
Juglandaceae	*Carys illinoinensis*																														A										
Juglandaceae	*Juglans regia*									A														A							A										A	
Lamiaceae	*Ajuga reptans*													Y																											
Lamiaceae	*Caryopteris incana*																							Y																		
Lamiaceae	*Clerodendrum thomsoniae*													Y																											
Lamiaceae	*Coleus blumei*													Y																											Y	

Family	Species	ArMV	AVA	AARSV	CsAlV	GDeFV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Lamiaceae	*Glechoma hederacea*	A																																								
Lamiaceae	*Lamium amplexicaule*	Y													A									A																			
Lamiaceae	*Lamium hybridum*																							Yt																			
Lamiaceae	*Lamium purpureum*																							Yt																			
Lamiaceae	*Mentha arvensis*	Y																																								
Lamiaceae	*Mentha piperita*													Nx																												
Lamiaceae	*Mentha spicata*													At																												
Lamiaceae	*Mentha spp.*													A																												
Lamiaceae	*Ocymum basilicum*	Y	N			Nc						Y			Nx					Y					Y						N	Nx		N								
Lamiaceae	*Ocymum canum*														Nx																												
Lamiaceae	*Physostegia virginia*													Y																												Nx
Lamiaceae	*Salvia splendens*													Y										Y																		A
Lamiaceae	*Syringa vulgaris*	Y																														A										
Lilliaceae	*Erythronium americanum*													A																												
Lilliaceae	*Lilium lancifolium*	A																																								
Lilliaceae	*Lilium. spp.*	Y												Y																											A	
Lilliaceae	*Llium longiflora*													Y																												
Lilliaceae	*Tulipa gesneriama*																	A						A																		
Lilliaceae	*Tulipa spp.*	Y												Y				A																								
Linaceae	*Linum usitatissimum*													Nc																											Nc	
Linderniaceae	*Torenia fournieri*	Y												Y	Y									Y								Y									Nc	
Lythraceae	*Lythrum salicaria*													Nc																											Nc	
Lythraceae	*Punica granatum*																															A									A	
Malvaceae	*Abelmoschus esculentus / Hibiscus esculentus*													Y								N																			Nc	
Malvaceae	*Abelmoschus manihot*													Nc																											Y	
Malvaceae	*Abutilon theophrasti*													Y																												
Malvaceae	*Alcea rosea / Althaea rosea*													Y										N																		
Malvaceae	*Althaea officinalis*																																								A	
Malvaceae	*Corchorus trilocularis*																																								A	
Malvaceae	*Gossypium hirsutum*													A	Nx																												
Malvaceae	*Hibiscus cannabinus*																																				Ym						
Malvaceae	*Hibiscus rosa-sinensis*																																				Y						
Malvaceae	*Lavatera trimestris*													Y																		Y										
Malvaceae	*Malva sp.*																									Y																
Malvaceae	*Malvastrum coromandelianum*													Y																												
Malvaceae	*Theobroma cacao*																	Y																								
Martyniaceae	*Proboscidea louisianica*													Y																											Y	
Monitiaceae	*Monitia perfoliata*		Y																																					Yx	
Moraceae	*Morus alba*												Y										Yt																		
Nyctaginaceae	*Boerhavia diffusa*																																								A	
Nyctaginaceae	*Boerhavia erecta*													Y																												
Nyctaginaceae	*Mirabilis jalapa*													Y												Y															Y	
Oleaceae	*Abeliophyllum distichum*																							A																		
Oleaceae	*Forestiera acuminata*																							A																		
Oleaceae	*Forsythia × intermedia*	Y																						A																		
Oleaceae	*Forsythia ovata*													Y																												

Family	Species	AfMV	AVA	AARSV	C_{sA}IV	GDeFV	GFLV	OLRSV	PBRSV	R_{pR}SV	MMMoV	AeRSV	MMLR$_{aV}$	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	C_{aw}YV	PRMV	PVU	ToRSV	SLSV
Oleaceae	*Forsythia sp.*									Y														A																		
Oleaceae	*Fraxinus americana*	A												A																											A	
Oleaceae	*Fraxinus excelsior*	Y																													Y											
Oleaceae	*Fraxinus spp.*													A										A							Y											
Oleaceae	*Jasminum mesnyi*	A																																								
Oleaceae	*Jasminum officinale*	Y												A																												
Oleaceae	*Ligustrum ovalifolium*	A																																								
Oleaceae	*Ligustrum sp.*																														Y											
Oleaceae	*Ligustrum vulgare*	Y								Y														A																		
Oleaceae	*Olea europaea*	A						Y																							A											
Oleaceae	*Syringia vulgaris*	A																																								
Onagraceae	*Clarkia amoena*													Y																		Y										
Onagraceae	*Clarkia elegans*													Y																		Y										
Orchidaceae	*Cymbidium sp.*																							Y								A										
Orchidaceae	*Ponthieva racemosa*																															A										
Osmundaceae	*Osmunda cinnamomea*													A																												
Oxalidaceae	*Oxalis repens*													Nx																												
Oxalidaceae	*Oxalis spp.*																							Y																		
Oxalidaceae	*Oxalis triangularis*															A																										
Oxalidaceae	*Oxalis tuberosa*								A																																	
Paeoniaceae	*Paeonia lactiflora*																		A																							
Paeoniaceae	*Paeonia suffruticosa*																		A																							
Papaveraceae	*Dicentra formosa*	A																								A																
Papaveraceae	*Dicentra spectabilis*													A																												
Papaveraceae	*Eschscholzia californica*													Y												Nc																
Papaveraceae	*Fumaria officinalis*																							Y																		
Papaveraceae	*Lamprocapnos spectabilis*													Y										A																		
Papaveraceae	*Papaver orientale*														Y											Nc																
Papaveraceae	*Papaver rhoeas*														Y										Y																	
Paulowniaceae	*Paulownia tomentosa*														Y																										
Pedaliaceae	*Sesamum indicum*		Y											Y					Y			Y									Y											
Phrymaceae	*Erythranthe moschata*													Y																										Y	
Phytolaccaceae	*Phytolacca americana / Phytolacca decanda*											Y		Y										Y																Y	
Piceaceae	*Picea sitchensis*																							A																	
Plantaginaceae	*Antirrhinum majus*	Nc												Y	Nx									Yx							Yx								Nx	Y		
Plantaginaceae	*Cymbalaria muralis*													Y																										Y		
Plantaginaceae	*Digitalis lanata*														Nx																											
Plantaginaceae	*Digitalis spp.*													Nc																												
Plantaginaceae	*Penstemon murrayanus*																																								A	
Plantaginaceae	*Penstemon sp.*													Y																											Y	
Plantaginaceae	*Plantago lanceolata*	Y												Y																N											
Plantaginaceae	*Plantago major*													A																N											
Plantaginaceae	*Plantago rugelii*																																								
Plantaginaceae	*Plantago spp.*																															A									
Plantaginaceae	*Veronica agrestis*									A																					Yx										
Plantaginaceae	*Veronica longifolia*													Y										A																Y	
Plantaginaceae	*Veronica officinalis*	A																																							

Family	Species	ArMV	AVA	AARSV	C_sAIV	GDeFV	GFLV	OLRSV	PBRSV	R_pRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	$C_{aw}YV$	PRMV	PVU	ToRSV	SLSV
Plantaginaceae	*Veronica persica*									A														A																		
Plumbaginaceae	*Limonium sinuatum*													Nc																												
Plumbaginaceae	*Statice armeria*													Nc																											Nc	
Poaceae	*Avena sativa*	Nc														Yt								N																		
Poaceae	*Chloris gayana*																							N																		
Poaceae	*Cynodon dactylon*						A																																			
Poaceae	*Dactylis glomerata*													Nx																												
Poaceae	*Echinochloa crus-galli*													Y																												
Poaceae	*Elymus repens*													Y																												
Poaceae	*Hordeum vulgare*	Y																										N												Nc	
Poaceae	*Lolium perenne*																							A																		
Poaceae	*Lolium pratense*													Y																												
Poaceae	*Phleum pratense*													Nx																												
Poaceae	*Poa annua*	Ys																																								
Poaceae	*Secale cereale*													Nc																											Nc	
Poaceae	*Sorghum bicolor*																																								Nc	
Poaceae	*Triticum aestivum*													Nc		Yt	N											N														Nc	
Poaceae	*Zea mays*								N					Y						N			N					N													Nc	
Polemoniaceae	*Gilia capitata*													Y																											Nc	
Polemoniaceae	*Gilia liniflora*													Y																											Y	
Polemoniaceae	*Phlox drummondii*	Y								A				Bc	Y					N													Yx								Y	
Polemoniaceae	*Phlox paniculata*	A																		Yt																						
Polemoniaceae	*Phlox spp.*	Y								Y										A																						
Polemoniaceae	*Phlox stolonifera*													Y																											A	
Polemoniaceae	*Phlox subulata*													Y						Nc				Yo																	Nc	
Polygonaceae	*Fagopyrum esculentum*	Y												Y										Yt	Y																	
Polygonaceae	*Fallopia convolvulus*									A																																
Polygonaceae	*Persicaria lapathifolia*													Y										Yt																		
Polygonaceae	*Persicaria maculosa / Polygonum persicaria*	Y																																								
Polygonaceae	*Polygonum aviculare*	Y												Y										A																		
Polygonaceae	*Polygonum hydropiper*													Y																												
Polygonaceae	*Rheum rhabarbarum*	A																						A					Y											A	
Polygonaceae	*Rheum rhaponticum*	Y												Y															A											Y	
Polygonaceae	*Rheum spp.*	A												Y															Y												
Polygonaceae	*Rheum × hybridum*																												Y											A	
Polygonaceae	*Rumex acetosella*													A										Y																A	
Polygonaceae	*Rumex crispus*																							Y																Y	
Polygonaceae	*Rumex obtusifolius*													Nx																	A										
Polygonaceae	*Rumex patienta*													Y																											Nc	
Portulaceae	*Portulaca grandiflora*																																								A	
Portulaceae	*Portulaca oleraceae*																																									
Primulaceae	*Atagallis arvensis*	Y																																								
Primulaceae	*Lycimachia nummularia*													A																												
Primulaceae	*Primula obconica*													Y										Yx																	Nc	
Primulaceae	*Primula sieboldii*																			A																						
Primulaceae	*Primula × pubescens*																							Nc																		
Proteaceae	*Grevillea robusta*													Nc																											Nc	

Family	Species	ArMV	AVA	AARSV	CsAlV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Ranunculaceae	*Aquilegia caerulea*													Y																												Nc	
Ranunculaceae	*Aquilegia spp.*																							N																			
Ranunculaceae	*Aquilegia vulgaris*																							Nc																		A	
Ranunculaceae	*Atemone coronaria*													A																													
Ranunculaceae	*Atemone spp.*													A																													
Ranunculaceae	*Clematis sp.*																							A																			
Ranunculaceae	*Delphinium elatum*									A																						A											
Ranunculaceae	*Delphinium hybridum*	Y												Nc																		Y										Nc	
Ranunculaceae	*Helleborus foetidus*																																									A	
Ranunculaceae	*Pulastilla vulgaris*																							Ye																			
Ranunculaceae	*Ranunculus asiaticus*																																Nx										
Ranunculaceae	*Ranunculus repens*	A																						Y																			
Ranunculaceae	*Ranunculus acris*																							Ye																			
Resedaceae	*Reseda alba*																								Y																		
Resedaceae	*Reseda odorata*													Nc																												Y	
Rhaminaceae	*Frangula alnus*																														Y											
Rosaceae	*Aruncus dioicus*																							Yo																			
Rosaceae	*Crataegus sp.*													A																													
Rosaceae	*Cydonia oblonga*																														A											
Rosaceae	*Fragaria × ananassa*	At						Y		A				A										A																		A	
Rosaceae	*Fragaria chiloensis*																																									Y	
Rosaceae	*Fragaria vesca*	Y								A					Yt									A																		A	
Rosaceae	*Fragaria virginiAta*																																									Y	
Rosaceae	*Geum quellyon*													Nc																												Nc	
Rosaceae	*Malus domestica / Malus pumila*															A								A								Y										At	
Rosaceae	*Malus orientalis*																																									At	
Rosaceae	*Malus sylvestris*																																									Y	
Rosaceae	*Potentilla sp.*													A																													
Rosaceae	*Prunus americana × P. salicina*						N																																				
Rosaceae	*Prunus amygdalus / Prunus dulcis*																							A																		A	
Rosaceae	*Prunus armeniaca*						N							Y																													
Rosaceae	*Prunus armeniana / Prunus hortulana*																																							Y		Y	
Rosaceae	*Prunus angustifolia*																																							A			
Rosaceae	*Prunus avium*	Y						Y		A				A										A								A						A				A	
Rosaceae	*Prunus besseyi*																																					Y					
Rosaceae	*Prunus cerasifera*						N																									Y						Y				Y	
Rosaceae	*Prunus cerasus*	Y						Y																Y		Y						A											
Rosaceae	*Prunus davidina*																																					Y					
Rosaceae	*Prunus domestica*	Y					N																									A						Y				Y	
Rosaceae	*Prunus dulcis*																																									A	
Rosaceae	*Prunus injucunda*																															A											
Rosaceae	*Prunus institia*																																					Y					
Rosaceae	*Prunus mahaleb*	Y					N	Y																Y																			
Rosaceae	*Prunus pennsylvatica*																															Y											

Family	Species	ArMV	AVA	AARSV	C_sAIV	G_DeFV	GFLV	OLRSV	PBRSV	R_pRSV	$MMLO_V$	Ae_RSV	$MMLR_aV$	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	$BLMo_V$	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	$C_{aw}YV$	PRMV	PVU	ToRSV	SLSV
Rosaceae	*Prunus persica*	Y					N			A						Yt								A							A						A		Yt		At	
Rosaceae	*Prunus salicina*	A																																			A					
Rosaceae	*Prunus serotina*													A																											
Rosaceae	*Prunus serrulata*						N																																		
Rosaceae	*Prunus spinosa*																														A										
Rosaceae	*Prunus spp.*	A								A														A							A										
Rosaceae	*Prunus tomentosa*																																							A	
Rosaceae	*Rosa canina*	A																																						A	
Rosaceae	*Rosa chinensis*	A																																							
Rosaceae	*Rosa hybrida*									A																															
Rosaceae	*Rosa indica*	A																																							
Rosaceae	*Rosa multiflora*	A																																							
Rosaceae	*Rosa spp.*	Y												A																	A									A	
Rosaceae	*Rosa × damascena*	A																																							
Rosaceae	*Rubus allegheniensis*													A																											
Rosaceae	*Rubus argutus*													A																											
Rosaceae	*Rubus flagellaris*													A																											
Rosaceae	*Rubus fruticosus*	A						Y						A										A							Y									A	
Rosaceae	*Rubus idaeus*	Yt					N			A				A										A							A									At	
Rosaceae	*Rubus lacinatus*																																							At	
Rosaceae	*Rubus obtusifolius*													A																											
Rosaceae	*Rubus occidentalis*						N							A																	Y									Y	
Rosaceae	*Rubus procerus*	Y																						A							A									Y	
Rosaceae	*Rubus spp.*							Y																A							A										
Rosaceae	*Sorbus acuparia*																														Y										
Rosaceae	*Spiraea douglasii*	A																																							
Rubiaceae	*Galium odoratum / Asperula odorata*													Nc																										Y	
Rubiaceae	*Mitracarpus villosus*																																							A	
Rutaceae	*Citrus trifoliata*						N																																		
Rutaceae	*Ptelea trifoliata*																														A										
Salicaceae	*Populus spp.*	A						Y						A										Y																	
Salicaceae	*Populus tremuloides*													Y																											
Salicaceae	*Salix spp.*	Y						Y						A																											
Sapindaceae	*Acer rubrum*																																						A		
Sapindaceae	*Cardiospermum halicacabum*													Y																											Nc	
Saxifragaceae	*Astilbe chinensis*													A																											
Saxifragaceae	*Astilbe sp.*	Y																						Y																	
Saxifragaceae	*Bergenia x schmidtii*																							Nc																	
Saxifragaceae	*Heuchera sanguinea*													Nc																											Nc	
Saxifragaceae	*Saxifraga sp.*													Nc																											Nc	
Scrophulariaceae	*Buddleja davidii*																																							A	
Scrophulariaceae	*Linaria macedonica*													Y																											Y	
Scrophulariaceae	*Nemesia strumosa*													Y																											Y	
Scrophulariaceae	*Verbascum phoeniceum*													Y																											Y	
Scrophulariaceae	*Verbascum thapsus*																																							A	
Scrophulariaceae	*Zaluzianskya villosa*													Y																											Y	
Solanaceae	*Atropa belladonna*																							Yx																	

Family	Species	ArMV	AVA	AARSV	CsAlV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Solanaceae	*Browallia speciosa*																													Y											
Solanaceae	*Browallia viscosa*																													Y											
Solanaceae	*Brugmansia spp.*													Y										Y																	
Solanaceae	*Calibrachoa spp.*													Y										Y																	
Solanaceae	*Capsicum annuum*		N									Nx		Y	Nx		N					N			Nc							Nx			N					Bl
Solanaceae	*Capsicum baccatum var. pendulum*								N																																
Solanaceae	*Capsicum frutescens*	Nc												Y	Y									Nc						Y											Nc
Solanaceae	*Capsicum sinensis*																																						N		
Solanaceae	*Capsicum spp.*													A										Nc																	A
Solanaceae	*Datura metel*									Y		Y		Y	Y															N	Nx							Nc	Y	Nc
Solanaceae	*Datura stramonium*	Nc	Y			Nc			Y	Y	Y	Y		Yl			N		Nc	Yx	Y			Y	Yl		Yx	Y	N		N	Y	Nx			Y			Y	Y	B
Solanaceae	*Datura tatula*								Y																						N										
Solanaceae	*Hyoscamus alba*													Y																											
Solanaceae	*Hyoscyamus niger*		Y											Y	Nx		N							Y																Y	N
Solanaceae	*Jaltomata procumbens*													Y																											
Solanaceae	*Lycium chinense*																		Nc																							
Solanaceae	*Niccandra physaloides*	Y	Y						Ym	Y				Y	Nx									Y																Y	Y
Solanaceae	*Nicotiana acuminata*													Y	Y																											
Solanaceae	*Nicotiana affinis*																							Y																	
Solanaceae	*Nicotiana alata*														Y															Y											
Solanaceae	*Nicotiana angustifolia*													Y										Y																	
Solanaceae	*Nicotiana attenuata*						Y																																		
Solanaceae	*Nicotiana benthamiana*				Yp	Yx	Ym		Y		Ym	Yp								Nc	Y			Yx	Y	Y	Y		Y	Y	Y	Y	Nx	Yx				Y			Y
Solanaceae	*Nicotiana bigelovii*		Y				Y		Y					Y																									Y	Y
Solanaceae	*Nicotiana cavicola*					Yx														Y														Yl						Y
Solanaceae	*Nicotiana clevelandii*	Yp	Yp		Yp	Yx	Ym	Y	Y	Yp	Ym			Ypl	Y		Y	N	Bx	Y	Y	Y		Yp	Y	Y	Y	Ym	Ym		Yp	Y	Y	Y	Yp	Y	Y			Y	Yp
Solanaceae	*Nicotiana debneyi*		Y				N		Y						Y										Y		Y	Y	Y		Y								Nc	Y	
Solanaceae	*Nicotiana fragrans*														Y																										
Solanaceae	*Nicotiana glauca*									Y														Y																	
Solanaceae	*Nicotiana glutinosa*	Nc	Y			Yx	Y	N	Y		Y	Y		Ym			Y	N	Bx	Y	Y	N		Y	Y		Yx	Y	Y			Y		N		Y			Y	Yx	B
Solanaceae	*Nicotiana langsdorfii*																							Y								Nx								
Solanaceae	*Nicotiana longiflora*													Y	Y																										
Solanaceae	*Nicotiana megalosiphon*						Y		Y						Y																Ys		Ym	Y							Y
Solanaceae	*Nicotiana multivalvis*																															Y									
Solanaceae	*Nicotiana occidentalis*		Y			Y			Y			Y			Y						Y			Y		Y	Y				Y			Y					Y	Y
Solanaceae	*Nicotiana panicullata*													Y										Y																	
Solanaceae	*Nicotiana plumbaginifolia*													Y																											
Solanaceae	*Nicotiana quadrivalva*													Y																											
Solanaceae	*Nicotiana repanda*													Y																											
Solanaceae	*Nicotiana rotundifolia*																																		Yl						
Solanaceae	*Nicotiana rustica*	Y					N		Y	Ym		Y		Y	Y				Bx					Ym	Y		Yx	Y	Y		Ym	Y	Yx	Yl		Y				Yx	Y
Solanaceae	*Nicotiana sanderae*													Y																											
Solanaceae	*Nicotiana spp.*																																								Ym
Solanaceae	*Nicotiana suaveolens*													Y																											
Solanaceae	*Nicotiana sylvestris*	Y										Y		Y										Y					Y										Y		Y
Solanaceae	*Nicotiana tabacum*	Y	Y		N		N	N	Y	Y	Y	Y		Yml	Yt	Yt	Y	N	B		Y	N		Y	Y	Y	Yx	Y	Y		Yml	Y	Yx	Y		Y	Y		Nx	Y	Yl
Solanaceae	*Nicotiana tomentosa*													Y																											

Family	Species	ArMV	AVA	AARSV	CsAIV	GDeFV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRS	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Solanaceae	*Nicotiana trigonophylla*													Yl																												
Solanaceae	*Petunia × hybrida*	Ym	Y				Y		Y	Yp	Y			Yl		Yt	Y	N	B		Y			Yp	Y		Bx	Y	N	Y	Yx	Y		Y						Nx	Y	Yp
Solanaceae	*Petunia spp.*													A								N		Yp																		
Solanaceae	*Petunia violacea*								Y					Y																												
Solanaceae	*Petuna nyctaginiflora*														Y																											
Solanaceae	*Physalis angulata*													Y																												
Solanaceae	*Physalis floridana*	Y	Y						Y					Y	Y		N		B					Y	Y		Nc				Yx					N				Y	A	
Solanaceae	*Physalis peruviana*		Y						Y					Y			N	N	B												Y									Y	Nc	
Solanaceae	*Salpiglossa sinuata*													Y																												
Solanaceae	*Schizanthus spp.*																							Yx																		
Solanaceae	*Solanum acaule*																															A										
Solanaceae	*Solanum americanum / nodiflorum*	Bl	Y											Y																												
Solanaceae	*Solanum bertholtii*																																									
Solanaceae	*Solanum betaceum / Cyphomandra betacea*	At													Nx																											
Solanaceae	*Solanum carolinense*													Y																												
Solanaceae	*Solanum chacoense*		N																													A										
Solanaceae	*Solanum chanceyense*								Y																																	
Solanaceae	*Solanum chinense*																																									
Solanaceae	*Solanum demissum*		Y																								Y															
Solanaceae	*Solanum demissum x S. tuberosum*								Y																																	
Solanaceae	*Solanum diphyllum*													Y																												
Solanaceae	*Solanum dulcamara*													Nx																												
Solanaceae	*Solanum elaeagnifolium*													Y																												
Solanaceae	*Solanum goniocalyx*																																									
Solanaceae	*Solanum lycopersicum / Lycopersicon esculentum*	Bl						Y	B	Y	Y	Y		Bl			N	N	Nc			N		B	Y		Yx	Y	N	Y	N	Y					N			Yx	B	
Solanaceae	*Solanum megistracrolobum*																																							Y		
Solanaceae	*Solanum melongena*		N											Yt	Y							N			Nc			Nx								N				Y		
Solanaceae	*Solanum microdontum*																																							Y		
Solanaceae	*Solanum mochicense*																					N																		Y		
Solanaceae	*Solanum nigrum*	Y							Y					Y	Nx																											Y	
Solanaceae	*Solanum pimpinellifolium / Lycopersicon pimpinellifolium*								Y						Y									N																Y	Y	
Solanaceae	*Solanum pseudo-capsicum*													Y																		N										
Solanaceae	*Solanum rostratum*																															Y										
Solanaceae	*Solanum sisymbriifolium*																																							Y		
Solanaceae	*Solanum stoloniferum*																																									
Solanaceae	*Solanum tuberosum*	Y	Y						Y					Bc	Nx	Yt								A	Nc			Y												Y	Y	
Staphylaceae	*Staphylea trifolia*	A								Y															A																	
Thymelaeaceae	*Daphne mezereum*	Y																	A																					Y	A	
Thymelaeaceae	*Daphne odora*													Y					A																							
Thymelaeaceae	*Daphne odora*																		A																							
Thymelaeaceae	*Daphne sp.*																																									
Thymelaeaceae	*Tropaelum peregrinum*													Y																												Nc
Thymelaeaceae	*Tropaeolum majus*	Y												Y											Y			Y	Nx												Nc

Family	*Species*	ArMV	AVA	AARSV	CsAIV	GDefV	GFLV	OLRSV	PBRSV	RpRSV	MMMoV	AeRSV	MMLRaV	TRSV	AILV	BRSV	CNV	CCLV	CNSV	GARSV	GCMV	MRSV	PVB	TBRV	AYRSV	ALRSV	BRV	CGMV	BLMoV	BLSV	CLRV	CYMV	GBLV	GTRSV	HLRSV	LALV	MLRSV	CawYV	PRMV	PVU	ToRSV	SLSV
Ulmaceae	*Ulmus americana*													A																A												Y	
Ulmaceae	*Ulmus sp.*	A																																									
Urticaceae	*Parietaria officinalis*																								Y																		
Urticaceae	*Urtica dioica*	Y																						Nc	Y																		
Urticaceae	*Urtica urens*	Y														A								Y																		
Verbeneaceae	*Lantanta camara*													Nc																													Nc	
Verbeneaceae	*Verbena rigida*													Y																													Y	
Violaceae	*Viola arvensis*																							Y																		
Violaceae	*Viola cornuta*													Y																													Y	
Violaceae	*Viola papilioNxcea*													Y																														
Violaceae	*Viola sp.*																							A																				
Violaceae	*Viola tricolor*													Y										Yx																				
Vitacaeae	*Vitis aestivalis*													A																														
Vitacaeae	*Vitis labrusca*						Y																						A				A								A	
Vitacaeae	*Vitis rupestris*						Y																										Y									
Vitacaeae	*Vitis sp.*	A					Y	Y						A	Y					Y				A																			
Vitacaeae	*Vitis vinifera*	Y				Y	Yt			A				At							Y			A							A		A	Y					Y		At	
Zamiaceae	*Zamia furfuraceae*													A																													
	Subgroup	A													B										C																		

References

Abad, J. A., Moyer, J. W., Kennedy, G. G., Holmes, G. A., and Cubeta, M. A. (2005). Tomato spotted wilt virus on potato in eastern North Carolina. *American Journal of Potato Research*, 82(3):255–261.

Abraham, P., Banwo, O. O., Kashina, B. D., and Alegbejo, M. D. (2021). Detection of weed species infected by tomato ring spot virus in field-grown tomato in Sudan savanna, Nigeria. *Nigerian Journal of Plant Protection*, 35(2):1–15.

Adams, I. and Fox, A. (2016). *Current Research Topics in Plant Virology*, chapter Diagnosis of Plant Viruses Using Next-Generation Sequencing and Metagenomic Analysis, pages 323–335. Springer International Publishing.

Adams, I. P., Boonham, N., and Jones, R. A. C. (2017). First complete genome sequence of *Arracacha virus A* isolated from a 38-year-old sample from Peru. *Genome Announcements*, 5(18):e00141–17.

Adams, I. P., Boonham, N., and Jones, R. A. C. (2018a). A 33-year-old plant sample contributes the first complete genomic sequence of *Potato Virus U*. *Microbiology Resource Announcements*, 7(23).

Adams, I. P., Boonham, N., Jones, R. A. C., and Stajich, J. (2018b). Full-genome sequencing of a virus from a 33-year-old sample demonstrates that *Arracacha mottle virus* is synonymous with *Arracacha virus Y*. *Microbiology Resource Announcements*, 7(21):e01393–18.

Adams, I. P., Fox, A., Boonham, N., and Jones, R. A. C. (2018c). Complete genomic sequence of the potyvirus *Mashua virus Y*, obtained from a 33-year-old mashua (*Tropaeaolum tuberosum*) sample. *Microbiology Resource Announcements*, 7:15.

Adams, I. P., Fox, A., Boonham, N., Massart, S., and De Jonghe, K. (2018d). The impact of high throughput sequencing on plant health diagnostics. *European Journal of Plant Pathology*, 152:909–919.

Adams, I. P., Glover, R., Souza-Richards, R., Bennett, S., Hany, U., and Boonham, N. (2013). Complete genome sequence of arracacha virus B: a novel cheravirus. *Archives of Virology*, 158(4):909–913.

Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., and Boonham, N. (2009). Next-generation sequencing and metagenomic analysis: A universal diagnostic tool in plant virology. *Molecular Plant Pathology*, 10:537–545.

Adams, I. P., Skelton, A., Macarthur, R., Hodges, T., Hinds, H., Flint, L., Nath, P. D., Boonham, N., and Fox, A. (2014). *Carrot yellow leaf virus* is associated with carrot internal necrosis. *PLoS ONE*, 9(11):e109125.

Adams, M. J. and Antoniw, J. F. (2006). DPVweb: a comprehensive database of plant and fungal virus genes and genomes. *Nucleic Acids Research*, 34(suppl_1):D382–D385.

Adams, M. J., Antoniw, J. F., and Beaudoin, F. (2005a). Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. *Molecular Plant Pathology*, 6(4):471–487.

Adams, M. J., Antoniw, J. F., and Fauquet, C. M. (2005b). Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Archives of Virology*, 150(3):459–479.

Adkins, S. and Rosskopf, E. N. (2002). Key west nightshade, a new experimental host for plant viruses. *Plant Disease*, 86(12):1310–1314.

Afonso, A. M., Guerra, R., Cavaco, A. M., Pinto, P., Andrade, A., Duarte, A., Power, D. M., and Marques, N. T. (2017). Identification of asymptomatic plants infected with *Citrus tristeza virus* from a time series of leaf spectral characteristics. *Computers and electronics in agriculture*, 141:340–350.

Agrios, G. N. (2005). *Plant pathology*. Elsevier.

Ahmed, A., Abd El Nagdi, W., and Youssef, M. (2007). Potential role of *Xiphinema americanum* and *Meloidogyne incognita* in transmission of *Peach rosette mosaic virus*. *Journal of Plant Production*, 32(3):1865–1876.

Aiewsakun, P. and Katzourakis, A. (2015). Endogenous viruses: Connecting recent and ancient viral evolution. *Virology*, 479-480:26–37.

Al Abdullah, A., El Beaino, T., Saponari, M., Hallak, H., and Digiaro, M. (2005). Preliminary evaluation of the status of olive-infecting viruses in Syria. *EPPO Bulletin*, 35(2):249–252.

Al Rwahnih, M., Daubert, S., Golino, D., and Rowhani, A. (2009). Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology*, 387(2):395–401.

Al-Shehbaz, I. A. (2010). A synopsis of the South American *Lepidium* (*Brassicaceae*). *Darwiniana*, 48(2):141–167.

Allen, W. R., Van Schagen, J. G., and Eveleigh, E. S. (1982). Transmission of peach rosette mosaic virus to peach, grape, and cucumber by *Longidorus diadecturus* obtained from diseased orchards in Ontario. *Canadian Journal of Plant Pathology*, 4(1):16–18.

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., and Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, 19(10):535–544.

Anderson, T. C., Nguyen, T.-A., Adams, J. K., Garrett, N. M., Bopp, C. A., Baker, J. B., McNeil, C., Torres, P., Ettestad, P. J., and Erdman, M. M. (2016). Multistate outbreak of human *Salmonella* Typhimurium infections linked to live poultry from agricultural feed stores and mail-order hatcheries, United States 2013. *One Health*, 2:144–149.

Arbizu, C., Huamán, Z., and Golmirzaie, A. (1997). *Biodiversity in Trust: Conservation and Use of Plant Genetic Resources in CGIAR Centres*, chapter Other Andean Roots and Tubers, pages 39–56. Cambridge University Press, Cambridge.

Arias, T., Riaño-Pachón, D. M., and Di Stilio, V. S. (2021). Genomic and transcriptomic resources for candidate gene discovery in the Ranunculids. *Applications in Plant Sciences*, 9(1):e11407.

Asjes, C. J. (1979). Viruses and virus diseases in Dutch bulbous irises (*Iris hollandica*) in the Netherlands. *Netherlands Journal of Plant Pathology*, 85(6):269–279.

Asjes, C. J. and Segers, L. C. (1983). Incidence and control of necrotic leaf mosaic caused by Arabis mosaic virus in *Lilium tigrinum splendens* in the Netherlands. *Journal of Phytopathology*, 106(2):115–126.

Ayazpour, K. (2014). *Alphabetic list of plant viruses and viroids reported from Iran*. Jahrom, Iran.

Azcona, J. S. (2019). *Cryptic viruses in Black-grass: Investigating their role in plant stress tolerance*. PhD thesis, Institute for Agri-food Research and Innovation, Newcastle University.

Baker, C. A. and Adkins, S. (2007). Tobacco ringspot virus found in the cardboard cycad (*Zamia furfuracea*) in Florida. *Plant Disease*, 91(1):112–112.

Banerjee, N., Wang, J.-Y., and Zaitlin, M. (1995). A single nucleotide change in the coat protein gene of tobacco mosaic virus is involved in the induction of severe chlorosis. *Virology*, 207(1):234–239.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., and Prjibelski, A. D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, 19(5):455–477.

Barker, R. (2022). Preventing the introduction and spread of yacon viruses. https://planthealthportal.defra.gov.uk/assets/factsheets/Yacon-Biosecurity-guidence-2022-v3.pdf, DEFRA. [Online; accessed 4-April-2024].

Bawden, F. C. and Pirie, N. W. (1950). Some effects of freezing in the leaf, and of citrate in vitro, on the infectivity of a tobacco necrosis virus. *Microbiology*, 4(3):482–492.

Bayer, C. and Appel, O. (2003). *Flowering Plants · Dicotyledons:* Malvales, Capparales *and Non-betalain* Caryophyllales, chapter *Tropaeolaceae*, pages 400–404.

Bejarano, E. R., Khashoggi, A., Witty, M., and Lichtenstein, C. (1996). Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. *Proceedings of the National Academy of Sciences*, 93(2):759–764.

Bejerman, N. and Debat, H. (2022). Exploring the tymovirales landscape through metatranscriptomics data. *Archives of Virology*, 167:1785–1803.

Bejerman, N., Dietzgen, R., and Debat, H. (2023). Novel tri-segmented rhabdoviruses: A data mining expedition unveils the cryptic diversity of cytorhabdoviruses. *Viruses*, 15(12):2402.

Bejerman, N., Dietzgen, R. G., and Debat, H. (2021). Illuminating the plant rhabdovirus landscape through metatranscriptomics data. *Viruses*, 13(7):1304.

Bejerman, N., Dietzgen, R. G., and Debat, H. (2022). Unlocking the hidden genetic diversity of varicosaviruses, the neglected plant rhabdoviruses. *Pathogens*, 11(10):1127.

Bejerman, N., Giolitti, F., De Breuil, S., and Lenardon, S. (2010). Molecular characterization of Sunflower chlorotic mottle virus: a member of a distinct species in the genus *Potyvirus*. *Archives of Virology*, 155(8):1331–1335.

Bellardi, M. G., Nanni, G., and Bertaccini, A. (2000). Old and new viruses of lily in Italy. In *ISHS Acta Horticulturae 568: X International Symposium on Virus Diseases of Ornamental Plants*, volume 568, pages 215–220. Acta Horticulturae.

Bentley, K. and Evans, D. J. (2018). Mechanisms and consequences of positive-strand RNA virus recombination. *Journal of General Virology*, 99(10):1345–1356.

Bernardo, P., Golden, M., Akram, M., Naimuddin, Nadarajan, N., Fernandez, E., Granier, M., Rebelo, A. G., Peterschmitt, M., Martin, D. P., and Roumagnac, P. (2013). Identification and characterisation of a highly divergent geminivirus: evolutionary and taxonomic implications. *Virus Research*, 177(1):35–45.

Bernardo, P., Muhire, B., François, S., Deshoux, M., Hartnady, P., Farkas, K., Kraberger, S., Filloux, D., Fernandez, E., Galzi, S., Ferdinand, R., Granier, M., Marais, A., Monge Blasco, P., Candresse, T., Escriu, F., Varsani, A., Harkins, G. W., Martin, D. P., and Roumagnac, P. (2016). Molecular characterization and prevalence of two capulaviruses: Alfalfa leaf curl virus from France and Euphorbia caput-medusae latent virus from South Africa. *Virology*, 493:142–153.

Betancourth, C., Salomón, C., Montaño, S., Salazar, C., Uribe, P., Martínez, A., Muñoz, L., and Cuervo, M. (2020). First report of Sugarcane mosaic virus in achira (*Canna edulis* Ker.) in Nariño, Colombia. *Revista UDCA Actualidad & Divulgación Científica*, 23(1).

Bhat, A. I. and Rao, G. P. (2020a). *Characterization of Plant Viruses : Methods and Protocols*, chapter Development of Infectious Clone of Virus, pages 449–463. Springer US, New York, NY.

Bhat, A. I. and Rao, G. P. (2020b). *Springer Protocols Handbooks*, chapter Virus Transmission Through Pollen, pages 61–64. Springer US.

Biemond, P. C., Oguntade, O., Kumar, P. L., Stomph, T. J., Termorshuizen, A. J., and Struik, P. C. (2013). Does the informal seed system threaten cowpea seed health? *Crop Protection*, 43:166–174.

Blitvich, B. J., Beaty, B. J., Blair, C. D., Brault, A. C., Dobler, G., Drebot, M. A., Haddow, A. D., Kramer, L. D., LaBeaud, A. D., and Monath, T. P. (2018). Bunyavirus taxonomy: limitations and misconceptions associated with the current ICTV criteria used for species demarcation. *The American Journal of Tropical Medicine and Hygiene*, 99(1):11.

Blouin, A. G., Dubuis, N., Brodard, J., Apotheloz-Perret-gentil, L., Altenbach, D., and Schumpp, O. (2023). Symptomatic, widespread, and inconspicuous: new detection of tomato fruit blotch virus. *Phytopathologia Mediterranea*, 62:371–376.

Bock, K. (1994). Studies on cassava brown streak virus disease in Kenya. *Tropical Science*, 34(1):134–145.

Bonjean, A. P., Dequidt, C., and Sang, T. (2016). Rapeseed in China. *Oilseeds and Fats, Crops and Lipids*, 23(6):D605.

Bonnave, M., Bleeckx, T., Terrazas, F., and Bertin, P. (2016). Effect of the management of seed flows and mode of propagation on the genetic diversity in an Andean farming system: the case of oca (*Oxalis tuberosa* mol.). *Agriculture and Human Values*, 33(3):673–688.

Boonham, N., Walsh, K., Hims, M., Preston, S., North, J., and Barker, I. (2002). Biological and sequence comparisons of *Potato virus Y* isolates associated with potato tuber necrotic ringspot disease. *Plant Pathology*, 51(2):117–126.

Bos, L. (1981). Hundred years of Koch's Postulates and the history of etiology in plant virus research. *Netherlands Journal of Plant Pathology*, 87(3):91–110.

Bos, L. (1992). *Potyvirus Taxonomy*, volume 5 of *Archives of Virology. Supplementa*, chapter Potyviruses, chaos or order?, pages 31–46. Springer Vienna.

Botermans, M., Van de Vossenberg, B. T. L. H., Verhoeven, J. T. J., Roenhorst, J. W., Hooftman, M., Dekter, R., and Meekes, E. T. M. (2013). Development and validation of a real-time RT-PCR assay for generic detection of pospiviroids. *Journal of Virological Methods*, 187(1):43–50.

Boulton, M. I. (2008). *Plant Virology Protocols*, volume 451 of *Methods in Molecular Biology*, chapter Construction of infectious clones for DNA viruses: mastreviruses, pages 503–523. Humana Press.

Boylan-Pett, W., Ramsdell, D. C., Hoopingarner, R. A., and Hancock, J. F. (1991). Honeybee foraging behavior, in-hive survival of infectious, pollen-borne blueberry leaf mottle virus and transmission of the virus in highbush blueberry. *Phytopathology*, 81(11):1407–1412.

Bradbury, E. J. and Emshwiller, E. (2011). The role of organic acids in the domestication of *Oxalis tuberosa*: A new model for studying domestication resulting in opposing crop phenotypes. *Economic Botany*, 65(1):76–84.

Bragard, C., Dehnen-Schmutz, K., Di Serio, F., Gonthier, P., Jacques, M. A., Jaques Miret, J. A., Justesen, A. F., MacLeod, A., Magnusson, C. S., Milonas, P., Navas-Cortes, J. A., Parnell, S., Potting, R., Reignault, P. L., Thulke, H. H., Van der Werf, W., Vicent Civera, A., Zappalà, L., Lucchi, A., Urek, G., Gómez, P., Mosbach-Schulz, O., Maiorano, A., de la Peña, E., and Yuen, J. (2021). Commodity risk assessment of *Ullucus tuberosus* tubers from Peru. *EFSA Journal*, 19(3):e06428.

Bragard, C., Dehnen-Schmutz, K., Gonthier, P., Jacques, M., Jaques Miret, J. A., Justesen, A. F., Macleod, A., Magnusson, C. S., Milonas, P., Navas-Cortes, J. A., Parnell, S., Potting, R., Reignault, P. L., Thulke, H., Van Der Werf, W., Vicent Civera, A., Yuen, J., Zappalà, L., Candresse, T., Lacomme, C., Bottex, B., Oplaat, C., Roenhorst, A., Schenk, M., and Di Serio, F. (2020). Pest categorisation of non-EU viruses and viroids of potato. *EFSA Journal*, 18(1).

Brait, N., Hackl, T., Morel, C., Exbrayat, A., Gutierrez, S., and Lequime, S. (2024). A tale of caution: How endogenous viral elements affect virus discovery in transcriptomic data. *Virus Evolution*, 10(1):vead088.

Brako, L. and Zarucchi, J. L. (1993). *Catalogue of the flowering plants and gymnosperms of Peru*, volume 45 of *Monographs in Systemic Botany from the Missouri Botanical Garden*. Missouri Botanical Garden.

Brasier, C. (2005). Preventing invasive pathogens: deficiencies in the system.

Bratsch, S. A. and Lockhart, B. E. L. (2016). Preliminary characterization of a virus associated with *Euonymus alatus* dieback. In *ISHS Acta Horticulturae 1191: III International Symposium on Woody Ornamentals of the Temperate Zone*, pages 69–76.

Bridgmon, G. H. and Walker, J. C. (1952). *Gladiolus* as a virus reservoir. *Pythopathology*, 42(2):65–70.

Broadbent, A. (2009). Causation and models of disease in epidemiology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 40(4):302–311.

Brown, D. J. F., Kunz, P., Grunder, J., and Robertson, W. M. (1998). Differential transmission of cherry rosette nepovirus by populations of *Longidorus arthensis* (*Nematoda* : *Longidoridae*) with a description of the association of the virus with the odontostyle of its vector. *Fundamental and Applied Nematology*, 21(6):673–677.

Brunt, A. A. (1975). Viruses and virus diseases of irises in Britain. *Acta Horticulturae*, (No.47):45–50. Abstract only.

Brunt, A. A. (1996). *Viruses of plants : descriptions and lists from the VIDE database*. CAB International, Wallingford, Oxon, UK.

Brunt, A. A., Barton, R. J., Phillips, S., and Jones, R. A. C. (1982a). Ullucus virus C, a newly recognised comovirus infecting *Ullucus tuberosus* (*Basellaceae*). *Annals of Applied Biology*, 101(1):73–78.

Brunt, A. A., Barton, R. J., Phillips, S., and Lana, A. O. (1980). Hibiscus latent ringspot virus, a newly recognised virus from *Hibiscus rosa-sinensis* (*Malvaceae*) in western Nigeria. *Annals of Applied Biology*, 96(1):37–43. [As cited in VIDE database.].

Brunt, A. A., Crabtree, K., and Gibbs, A. (1990). *Viruses of tropical plants. Descriptions and lists from the VIDE database*. C.A.B. International, Wallingford.

Brunt, A. A., Phillips, S., Jones, R. A. C., and Kenten, R. H. (1982b). Viruses detected in *Ullucus tuberosus* (*Basellaceae*) from Peru and Bolivia. *Annals of Applied Biology*, 101(1):65–71.

Buchhop, J., von Bargen, S., and Büttner, C. (2009). Differentiation of *Cherry leaf roll virus* isolates from various host plants by immunocapture-reverse transcription-polymerase chain reaction-restriction fragment length polymorphism according to phylogenetic relations. *Journal of Virological Methods*, 157(2):147–154.

Buddenhagen, C. E., Hernandez Nopsa, J. F., Andersen, K. F., Andrade-Piedra, J., Forbes, G. A., Kromann, P., Thomas-Sharma, S., Useche, P., and Garrett, K. A. (2017). Epidemic network analysis for mitigation of invasive pathogens in seed systems: Potato in Ecuador. *Phytopathology*, 107(10):1209–1218.

Budzyńska, D., Minicka, J., Hasiów-Jaroszewska, B., and Elena, S. F. (2020). Molecular evolution of tomato black ring virus and de novo generation of a new type of defective RNAs during long-term passaging in different hosts. *Plant Pathology*, 69(9):1767–1776.

Budzyńska, D., Zwart, M. P., and Hasiów-Jaroszewska, B. (2022). Defective RNA particles of plant viruses—origin, structure and role in pathogenesis. *Viruses*, 14(12):2814.

Bujarski, J., Gallitelli, D., García-Arenal, F., Pallás, V., Palukaitis, P., Reddy, M. K., Wang, A., and Consortium, I. R. (2019). ICTV virus taxonomy profile: *Bromoviridae*. *Journal of General Virology*, 100(8):1206–1207.

Burgman, M., Roberts, B., Sansford, C., Griffin, R., and Mengersen, K. (2013). *The Handbook of Plant Biosecurity: Principles and Practices for the Identification, Containment and Control of Organisms that Threaten Agriculture and the Environment Globally*, chapter The role of pest risk analysis in plant biosecurity, pages 235–267. Springer.

Bustin, S. and Huggett, J. (2017). qPCR primer design revisited. *Biomolecular Detection and Quantification*, 14:19–28.

Bömer, M., Rathnayake, A. I., Visendi, P., Sewe, S. O., Sicat, J. P. A., Silva, G., Kumar, P. L., and Seal, S. E. (2019). Tissue culture and next-generation sequencing: A combined approach for detecting yam (*Dioscorea spp.*) viruses. *Physiological and Molecular Plant Pathology*, 105:54–66.

CABI (2024). CABI compendium. https://www.cabidigitallibrary.org/journal/cabicompendium. Wallingford, UK.

Cabinet Office (2023). UK biological security strategy. https://assets.publishing.service.gov.uk/media/64c0ded51e10bf000e17ceba/uk_biological_security_strategy.pdf. ISBN 978-1-5286-3761-9, CP 858.

Cabrera Mederos, D., Bejerman, N., Trucco, V., De Breuil, S., Lenardon, S., and Giolitti, F. (2017). Complete genome sequence of sunflower ring blotch virus, a new potyvirus infecting sunflower in Argentina. *Archives of Virology*, 162(6):1787–1790.

Cabrera Mederos, D., Debat, H., Torres, C., Portal, O., Jaramillo Zapata, M., Trucco, V., Flores, C., Ortiz, C., Badaracco, A., Acuña, L., Nome, C., Quito-Avila, D., Bejerman, N., Castellanos Collazo, O., Sánchez-Rodríguez, A., and Giolitti, F. (2022). An unwanted association: The threat to papaya crops by a novel potexvirus in Northwest Argentina. *Viruses*, 14(10):2297.

Caetano, B. F. R., De Moura, N. A., Almeida, A. P. S., Dias, M. C., Sivieri, K., and Barbisan, L. F. (2016). Yacon (*Smallanthus sonchifolius*) as a food supplement: Health-promoting benefits of fructooligosaccharides. *Nutrients*, 8(7):436.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*, 10:1–9.

Candresse, T., Filloux, D., Muhire, B., Julian, C., Galzi, S., Fort, G., Bernardo, P., Daugrois, J.-H., Fernandez, E., Martin, D. P., Varsani, A., and Roumagnac, P. (2014). Appearances can be deceptive: Revealing a hidden viral infection with deep sequencing in a plant quarantine context. *PLoS ONE*, 9(7):e102945.

Card, S. D., Pearson, M. N., and Clover, G. R. G. (2007). Plant pathogens transmitted by pollen. *Australasian Plant Pathology*, 36(5):455.

Castillo, R., Nieto, C., and Peralta, E. (1988). El germoplasma de cultivos andinos en Ecuador. In *VI Congreso Internacional sobre Cultivos Andinos*, pages 323–331. INIAP.

Castillo, R. O. (1995). Plant genetic resources in the Andes: Impact, conservation, and management. *Crop Science*, 35(2).

Castro, A., Caballero, M., Herbas, A., and Carballo, S. (2012). Antioxidants in yacon products and effect of long term storage. *Food Science and Technology*, 32:432–435.

Celli, M. G., Perotto, M. C., Luciani, C. E., Pozzi, E. A., and Conci, V. C. (2019). Molecular characterization of the garlic virus B genome and evidence of allexivirus recombination. *European journal of plant pathology*, 153:301–310.

CGIAR Technical Advisory Committee (1992). Diversity for development: the strategy of the international plant genetic resources institute (revised draft). https://hdl.handle.net/10568/118820. Internal report.

Chang, S. J., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, 11(2):113–116.

Chen, J., Shi, Y.-H., Adams, M. J., and Chen, J.-P. (2005). The complete sequence of the genomic RNA of an isolate of *Lily virus X* (genus *Potexvirus*). *Archives of Virology*, 150(4):825–832.

Chen, Q., Li, M., Wang, C., Li, Z., Xu, J., Zheng, Q., Liu, P., and Zhou, H. (2018). Combining targeted metabolites analysis and transcriptomics to reveal chemical composition difference and underlying transcriptional regulation in Maca (*Lepidium meyenii* Walp.) ecotypes. *Genes*, 9(7):335.

Chen, Y.-K., Chang, Y.-S., and Chen, C.-C. (2011). Identification of *Carnation mottle virus* from Lisianthus plants in Taiwan. *Plant Disease*, 95(8):1036–1036.

Cheng, C.-P. and Tsai, C.-H. (1999). Structural and functional analysis of the 3′ untranslated region of bamboo mosaic potexvirus genomic RNA. *Journal of Molecular Biology*, 288(4):555–565.

Chiapello, M., Rodríguez-Romero, J., Ayllón, M. A., and Turina, M. (2020). Analysis of the virome associated to grapevine downy mildew lesions reveals new mycovirus lineages. *Virus Evolution*, 6(2):veaa058.

Chirat, L. (1841). *Etude des fleurs : botanique élémentaire, descriptive et usuelle, simplifiée pour la jeunesse et les familles*, volume 2. Cormon et Blanc, Lyon, France.

Choi, H., Jo, Y., Chung, H., Choi, S. Y., Kim, S.-M., Hong, J.-S., Lee, B. C., and Cho, W. K. (2023). Investigating variability in viral presence and abundance across soybean seed development stages using transcriptome analysis. *Plants*, 12(18):3257.

Christie, S. R. and Crawford, W. E. (1978). Plant virus range of *Nicotiana benthamiana*. *Plant Disease Reporter*, 62:20–22.

Cigsar, I., Digiaro, M., Gokalp, K., Abou Ghanem-Sabanadzovic, N., De Stradis, A., Boscia, D., and Martell, G. P. (2003). Grapevine deformation virus, a novel nepovirus from Turkey. *Journal of Plant Pathology*, 85(3):183–191.

Clark, C. C. A., Crump, R., KilBride, A. L., and Green, L. E. (2016). Farm membership of voluntary welfare schemes results in better compliance with animal welfare legislation in Great Britain. *Animal Welfare*, 25(4):461–469.

Clark, M. F. and Adams, A. N. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of general virology*, 34(3):475–483.

Clark, T. A. (1999). Mashua. *Ethnobotanical Leaflets*, 1999(2):3.

Clarke, D. K., Duarte, E. A., Moya, A., Elena, S. F., Domingo, E., and Holland, J. (1993). Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *Journal of Virology*, 67(1):222–228.

Claverie, J.-M. (2020). Fundamental difficulties prevent the reconstruction of the deep phylogeny of viruses. *Viruses*, 12(10):1130.

Condori, B., Mamani, P., Botello, R., Patiño, F., Devaux, A., and Ledent, J. F. (2008). Agrophysiological characterisation and parametrisation of Andean tubers: Potato (*Solanum sp.*), oca (*Oxalis tuberosa*), isaño (*Tropaeolum tuberosum*) and papalisa (*Ullucus tuberosus*). *European Journal of Agronomy*, 28(4):526–540.

Constable, F., Kelly, G., and Dall, D. (2021). Viruses in cucurbit seeds from on-line mail-order providers. *Australasian Plant Disease Notes*, 16(1).

Converse, R. H. (1987). *Virus diseases of small fruits*, volume 631 of *Agricultural Handbook*. United States Department of Agriculture, Agricultural Research Service, Washington, DC.

Converse, R. H. and Ramsdell, D. C. (1982). Occurrence of tomato and tobacco ringspot viruses and of dagger and other nematodes associated with cultivated highbush blueberries in Oregon. *Plant disease*, 66(8):710–712.

Cooper, J. I. (1976). The possible epidemiological significance of pollen and seed transmission in the cherry leaf roll virus/*Betula spp.* complex. *Mitteilungen der Biologischen Bundesanstalt für Land-und Forstwirtschaft Berlin-Dahlem*, 170:17–27.

Cooper, J. I. and Sweet, J. B. (1976). The detection of viruses with nematode vectors in six woody hosts. *Forestry: An International Journal of Forest Research*, 49(1):73–78.

Cordero, M. A., Robbins, R. T., and Szalanski, A. L. (2012). Taxonomic and molecular identification of *Mesocriconema* and *Criconemoides* species (*Nematoda*: *Criconematidae*). *Journal of Nematology*, 44(4):399.

Coutts, B. A., Kehoe, M. A., and Jones, R. A. C. (2013). *Zucchini yellow mosaic virus*: Contact transmission, stability on surfaces, and inactivation with disinfectants. *Plant Disease*, 97(6):765–771.

Crespi, M., Messens, E., Caplan, A. B., Van Montagu, M., and Desomer, J. (1992). Fasciation induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. *The EMBO journal*, 11(3):795–804.

Crosslin, J. M., Eastwell, K. C., Davitt, C. M., and Abad, J. A. (2010). First report of seedborne *Cherry leaf roll virus* in wild potato, *Solanum acaule*, from South America. *Plant Disease*, 94(6):782–782.

Crosslin, J. M., Thomas, P. E., and Brown, C. R. (1999). Distribution of tobacco rattle virus in tubers of resistant and susceptible potatoes and systemic movement of virus into daughter plants. *American Journal of Potato Research*, 76(4):191–197.

Crowley, M. A. (2003). An introduction to the WTO and GATT. *Economic Perspectives*, 27(4):42–57.

Cunningham, S., Dobbie, S., Kell, S., Maxted, N., and Rosenfeld, A. (2018). The benefits of exotic food crops cultivated by small-scale growers in the UK. *Renewable Agriculture and Food Systems*, 33(6):569–584.

Daly, G. M., Leggett, R. M., Rowe, W., Stubbs, S., Wilkinson, M., Ramirez-Gonzalez, R. H., Caccamo, M., Bernal, W., and Heeney, J. L. (2015). Host subtraction, filtering and assembly validations for novel viral discovery using next generation sequencing data. *PLoS ONE*, 10(6):e0129059.

De Almeida Paula, H. A., Abranches, M. V., and de Luces Fortes Ferreira, C. L. (2015). Yacon (*Smallanthus sonchifolius*): A food with multiple functions. *Critical Reviews in Food Science and Nutrition*, 55(1):32–40.

De Bokx, J. A., Beemster, A., Lambers, R. H. R., Hiddema, J., Van Hoof, H. A., Maat, D. Z., Reestman, A. J., Rozendaal, A., Schepers, A., and Van Slogteren, D. M. (1972). *Viruses of potatoes and seed-potato production*. Centre for Agricultural Publishing and Documentation, Wageningen.

De Jong, H. (2016). Impact of the potato on society. *American Journal of Potato Research*, 93(5):415–429.

De Jonghe, K., De Krom, C., Fox, A., and Ziebell, H. (2022). Phytosanitary risks of newly introduced crops (PRONC). https://drop.euphresco.net/data/1d746fbd-9cea-4181-a89d-68dbdbe6b342https://drop.euphresco.net/getfile/3543a809-6fcc-4ae4-b5aa-9d4ff880ab43, EPPO Secretariat. Report.

De Souza, J., Fuentes, S., Müller, G., Gamarra, H., Guzmán, M., Cuellar, W., and Kreuze, J. (2021). High throughput sequencing for the detection and characterization of new virus found in arracacha (*Arracacia xanthorrhiza*). *Scientia Agropecuaria*, 12(4):471–480.

De Souza, J., Gamarra, H., Müller, G., and Kreuze, J. (2018). First report of Potato virus S naturally infecting arracacha (*Arracacia xanthorrhiza*) in Peru. *Plant Disease*, 102(2):460–460.

De Souza, J., Müller, G., Perez, W., Cuellar, W., and Kreuze, J. (2017). Complete sequence and variability of a new subgroup B nepovirus infecting potato in central Peru. *Archives of Virology*, 162(3):885–889.

Debat, H. and Bejerman, N. (2022). A glimpse into the DNA virome of the unique "living fossil" *Welwitschia mirabilis*. *Gene*, 843:146806.

Debat, H. and Bejerman, N. (2023). Two novel flavi-like viruses shed light on the plant-infecting koshoviruses. *Archives of Virology*, 168(7).

Debat, H., Garcia, M. L., and Bejerman, N. (2023). Expanding the repertoire of the plant-infecting ophioviruses through metatranscriptomics data. *Viruses*, 15(4):840.

Debrot, E. A. (1964). Studies on a strain of raspberry ringspot virus occurring in England. *Annals of Applied Biology*, 54(2):183–191.

DEFRA (2024). Prohibited commodities. https://planthealthportal.defra.gov.uk/trade/imports/imports-from/prohibitions/. [Online, accessed 4-April-2024.].

DEFRA and Forestry Commission, The Scottish Government, and Welsh Government (2023). Plant biosecurity strategy for Great Britain (2023 to 2028). https://www.gov.uk/government/publications/plant-biosecurity-strategy-for-great-britain-2023-to-2028/plant-biosecurity-strategy-for-great-britain-2023-to-2028. Government document.

Delfosse, V. C., Barrios Barón, M. P., and Distéfano, A. J. (2021). What we know about poleroviruses: Advances in understanding the functions of polerovirus proteins. *Plant Pathology*, 70(5):1047–1061.

Delhey, R., Monasterios, T., and Monasterio, T. (1977). A mosaic disease of isañu (*Tropaeolum tuberosum*) from Bolivia / Eine Mosaikkrankheit an Isañu (*Tropaeolum tuberosum*) aus Bolivien. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz / Journal of Plant Diseases and Protection*, 84(4):224–231.

Derraik, J. G. B. and Phillips, S. (2010). Online trade poses a threat to biosecurity in New Zealand. *Biological Invasions*, 12(6):1477–1480.

Desbiez, C. and Lecoq, H. (2004). The nucleotide sequence of *Watermelon mosaic virus* (WMV, *Potyvirus*) reveals interspecific recombination between two related potyviruses in the 5′ part of the genome. *Archives of Virology*, 149(8):1619–1632.

Desbiez, C., Moury, B., and Lecoq, H. (2011). The hallmarks of "green" viruses: Do plant viruses evolve differently from the others? *Infection, Genetics and Evolution*, 11(5):812–824.

Dieffenbach, C. W., Lowe, T. M., and Dveksler, G. S. (1993). General concepts for PCR primer design. *Genome Research*, 3(3):S30–S37.

Digiaro, M., Elbeaino, T., and Martelli, G. P. (2007). Development of degenerate and species-specific primers for the differential and simultaneous RT-PCR detection of grapevine-infecting nepoviruses of subgroups A, B and C. *Journal of Virological Methods*, 141(1):34–40.

Digiaro, M., Elbeaino, T., and Martelli, G. P. (2017). *Grapevine Viruses: Molecular Biology, Diagnostics and Management*, chapter *Grapevine fanleaf virus* and Other Old World Nepoviruses, pages 47–82. Springer International Publishing.

Dijkstra, J., De Jager, C. P., Dijkstra, J., and De Jager, C. P. (1998). Transmission by Thrips (*Thripidae*). *Practical Plant Virology: Protocols and Exercises*, pages 174–177.

Dolja, V. V., Boyko, V. P., Agranovsky, A. A., and Koonin, E. V. (1991). Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: Two families with distinct patterns of sequence and probably structure conservation. *Virology*, 184(1):79–86.

Domier, L. L., McCoppin, N. K., Larsen, R. C., and D'Arcy, C. J. (2002). Nucleotide sequence shows that *Bean leafroll virus* has a *Luteovirus*-like genome organization. *Journal of General Virology*, 83(7):1791–1798.

Domingo, E. (2002). Quasispecies theory in virology. *Journal of Virology*, 76(1):463–465.

Douthit, L. B. and McGuire, J. M. (1978). Transmission of tobacco ringspot virus by *Xiphinema americanum* to a range of hosts. *Plant Disease Reporter*, 62(2):164–166.

Dreyer, L. L., Esler, K. J., and Zietsman, J. (2006). Flowering phenology of South African *Oxalis*—possible indicator of climate change? *South African Journal of Botany*, 72(1):150–156.

Duarte, L., Alexandre, M., Rivas, E., Galleti, S., Harakava, R., and Chaves, A. (2014). A new Potyvirus species in *Cotyledon orbiculata* in mixed infection with a Nucleorhabdovirus. *Journal of Plant Pathology*, 96(1):143–149.

Duenas-Lopez, M. A. (2022). *Tropaeolum majus* (nasturtium). *CABI Compendium*.

Duffus, J. E. and Russell, G. E. (1970). Serological and host range evidence for the occurrence of beet western yellows virus in Europe. *Phytopathology*, 60(1):199–1202.

Duffy, S. and Seah, Y. M. (2010). 98% identical, 100% wrong: per cent nucleotide identity can lead plant virus epidemiology astray. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1548):1891–1897.

Dujovny, G., Usugi, T., Shohara, K., and Lenardon, S. (1998). Characterization of a potyvirus infecting sunflower in Argentina. *Plant Disease*, 82(5):470–474.

Dunleavy, J. M. (1957). The grasshopper as a vector of tobacco ringspot virus in soybean. *Phytopathology*, 47(11):681–682.

Dunn, M., Marzano, M., and Forster, J. (2020). Buying better biosecurity: Plant-buying behaviour and the implications for an accreditation scheme in the horticultural sector. *Plants, People, Planet*, 2(3):259–268.

Dupré, J. and Guttinger, S. (2016). Viruses as living processes. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 59:109–116.

Dupuis, B., Bragard, C., and Schumpp, O. (2019). Resistance of potato cultivars as a determinant factor of *Potato virus Y* (PVY) epidemiology. *Potato research*, 62:123–138.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5):1792–1797.

Edgar, R. C., Taylor, B., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J. F., De La Peña, M., Korobeynikov, A., Chikhi, R., and Babaian, A. (2022). Petabase-scale sequence alignment catalyses viral discovery. *Nature*, 602(7895):142–147.

EFSA Panel on Plant Health, . (2010). Guidance on a harmonised framework for pest risk assessment and the identification and evaluation of pest risk management options by EFSA. *EFSA Journal*, 8(2).

EFSA Panel on Plant Health, . (2013). Scientific opinion on the risk to plant health posed by *Arabis mosaic virus*, *Raspberry ringspot virus*, *Strawberry latent ringspot virus* and *Tomato black ring virus* to the EU territory with the identification and evaluation of risk reduction options. *EFSA Journal*, 11(10):3377.

Eid, S. and Pappu, H. R. (2014). Expression of endogenous para-retroviral genes and molecular analysis of the integration events in its plant host *Dahlia variabilis*. *Virus Genes*, 48(1):153–159.

Elbeaino, T., Digiaro, M., Ghebremeskel, S., and Martelli, G. P. (2012). *Grapevine deformation virus*: Completion of the sequence and evidence on its origin from recombination events between *Grapevine fanleaf virus* and *Arabis mosaic virus*. *Virus Research*, 166(1-2):136–140.

EPPO (2017). PM 7/2 (2) *Tobacco ringspot virus*. *EPPO Bulletin*, 47(2):135–145.

EPPO (2023a). EPPO A1 list of pests recommended for regulation as quarantine pests. https://www.eppo.int/ACTIVITIES/plant_quarantine/A1_list. [Online, accessed 27-February-2024.].

EPPO (2023b). EPPO A2 list of pests recommended for regulation as quarantine pests. https://www.eppo.int/ACTIVITIES/plant_quarantine/A2_list. [Online, accessed 27-February-2024.].

EPPO (2023c). EPPO Global Database. https://gd.eppo.int.

EPPO (2023d). PM 1/2(32): EPPO A1 and A2 lists of pests recommended for regulation as quarantine pests. Standard, EPPO.

Evans, A. S. (1976). Causation and disease: the Henle-Koch postulates revisited. *The Yale Journal of Biology and Medicine*, 49(2):175–195.

Evans, H. F., Richardson, B., and Hood, I. A. (2010). Pest risk analysis - organisms or pathways? *New Zealand Journal of Forestry Science*, 40.

Fabre, F., Montarry, J., Coville, J., Senoussi, R., Simon, V., and Moury, B. (2012). Modelling the evolutionary dynamics of viruses within their hosts: A case study using high-throughput sequencing. *PLoS Pathogens*, 8(4):e1002654.

Falkow, S. (1988). Molecular Koch's Postulates applied to microbial pathogenicity. *Reviews of Infectious Diseases*, 10:S274–S276.

Fan, W., Wang, S., Wang, H., Wang, A., Jiang, F., Liu, H., Zhao, H., Xu, D., and Zhang, Y. (2022). The genomes of chicory, endive, great burdock and yacon provide insights into *Asteraceae* palaeo-polyploidization history and plant inulin production. *Molecular Ecology Resources*, 22(8):3124–3140.

Feldman, J. M. and Gracia, O. (1977). Pepper severe mosaic virus: a new potyvirus from pepper in Argentina. *Phytopathologische Zeitschrift*, 89(2).

Ferguson, C. and Ali, A. (2024). Global genetic diversity of Tobacco ringspot virus including newly reported isolates from cotton (*Gossypium hirsutum* L.) in Oklahoma. *Plant Disease*, 108:635–646.

Fernández, E. C., Viehmannová, I., Lachman, J., and Milella, L. (2006). Yacon [*Smallanthus sonchifolius* (Poepping & endlicher) H. robinson]: a new crop in the central europe. *Plant Soil and Environment*, 52(12):564.

Ferrer, R. M., Ferriol, I., Moreno, P., Guerri, J., and Rubio, L. (2011). Genetic variation and evolutionary analysis of broad bean wilt virus 2. *Archives of virology*, 156:1445–1450.

Filardo, F., Nancarrow, N., Kehoe, M., Mctaggart, A. R., Congdon, B., Kumari, S., Aftab, M., Trębicki, P., Rodoni, B., Thomas, J., and Sharman, M. (2021). Genetic diversity and recombination between turnip yellows virus strains in Australia. *Archives of Virology*, 166(3):813–829.

Filloux, D., Dallot, S., Delaunay, A., Galzi, S., Jacquot, E., and Roumagnac, P. (2015a). *Plant Pathology*, chapter Metagenomics Approaches Based on Virion-Associated Nucleic Acids (VANA): An Innovative Tool for Assessing Without A Priori Viral Diversity of Plants, pages 249–257. Springer New York.

Filloux, D., Murrell, S., Koohapitagtam, M., Golden, M., Julian, C., Galzi, S., Uzest, M., Rodier-Goud, M., D'Hont, A., Vernerey, M. S., Wilkin, P., Peterschmitt, M., Winter, S., Murrell, B., Martin, D. P., and Roumagnac, P. (2015b). The genomes of many yam species contain transcriptionally active endogenous geminiviral sequences that may be functionally expressed. *Virus Evolution*, 1(1):vev002–vev002.

Fiore, N., Zamorano, A., Pino, A. M., González, F., Rosales, I. M., Sánchez-Navarro, J. A., and Pallás, V. (2016). Survey of stone fruit viruses and viroids in Chile. *Journal of Plant Pathology*, pages 631–635.

Fisher, J. R. (2013). First report of *Tobacco ringspot virus* infecting *Ajuga reptans* in ohio. *Plant Health Progress*, 14(1):56.

Fletcher, P. J. Fletcher, J. D. (2001). *In vitro* virus elimination in three Andean root crops: Oca (*Oxalis tuberosa*), ulluco (*Ullucus tuberosus*), and arracacha (*Arracacia xanthorrhiza*). *Plant Methods*, 29:23–27.

Fletcher, P. J., Fletcher, J. D., and Follett, J. M. (2005). Using tissue culture to help develop new crops. In *Combined Proceedings International Plant Propagators' Society*, volume 55, page 171. IPPS; 1998.

Flores, H. E., Walker, T. S., Guimaraes, R. L., Bais, H. P., and Vivanco, J. M. (2003). Andean root and tuber crops: Underground rainbows. *HortScience*, 38(2):161–167.

Fontdevila Pareta, N., Khalili, M., Maachi, A., Rivarez, M. P. S., Rollin, J., Salavert, F., Temple, C., Aranda, M. A., Boonham, N., Botermans, M., Candresse, T., Fox, A., Hernando, Y., Kutnjak, D., Marais, A., Petter, F., Ravnikar, M., Selmi, I., Tahzima, R., Trontin, C., Wetzel, T., and Massart, S. (2023). Managing the deluge of newly discovered plant viruses and viroids: an optimized scientific and regulatory framework for their characterization and risk analysis. *Frontiers in Microbiology*, 14.

Forbes, G. A., Charkowski, A., Andrade-Piedra, J., Parker, M. L., and Schulte-Geldermann, E. (2020). *The Potato Crop*, chapter Potato Seed Systems, pages 431–447. Springer International Publishing.

Forster, R. L. S. and Morris-Krsinich, B. A. M. (1985). A distinct strain of lucerne Australian latent virus in white clover in New Zealand. *Annals of Applied Biology*, 107(3):449–454.

Fowkes, A., Adams, I. P., Jones, R. A. C., Fox, A., Mcgreig, S., and Boonham, N. (2021a). Historical and recent tomato black ring virus and beet ringspot virus isolate genomes reveal interspecies recombination and plant health regulation inconsistencies. *Plant Pathology*. https://doi.org/10.1111/ppa.13507.

Fowkes, A. R., Mcgreig, S., Pufal, H., Duffy, S., Howard, B., Adams, I. P., Macarthur, R., Weekes, R., and Fox, A. (2021b). Integrating high throughput sequencing into survey design reveals turnip yellows virus and soybean dwarf virus in pea (*Pisum Sativum*) in the united kingdom. *Viruses*, 13(12):2530.

Fox, A. (2020). Reconsidering causal association in plant virology. *Plant Pathology*, 69(6):956–961.

Fox, A., Fowkes, A. R., Skelton, A., Harju, V., Buxton-Kirk, A., Kelly, M., Forde, S. M., Pufal, H., Conyers, C., Ward, R., Weekes, R., Boonham, N., and Adams, I. P. (2019). Using high-throughput sequencing in support of a plant health outbreak reveals novel viruses in *Ullucus tuberosus* (*Basellaceae*). *Plant Pathology*, 68:576–587.

Fox, A., Gibbs, A. J., Fowkes, A. R., Pufal, H., Mcgreig, S., Jones, R. A. C., Boonham, N., and Adams, I. P. (2022). Enhanced apiaceous potyvirus phylogeny, novel viruses, and new country and host records from sequencing *Apiaceae* samples. *Plants*, 11(15):1951.

Fox, A. and Mumford, R. A. (2017). Plant viruses and viroids in the United Kingdom: An analysis of first detections and novel discoveries from 1980 to 2014. *Virus Res*, 241:10–18.

Fraile, A., Escriu, F., Aranda, M. A., Malpica, J. M., Gibbs, A. J., and García-Arenal, F. (1997). A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *Journal of Virology*, 71(11):8316–8320.

Fraser, E. D. G. (2003). Social vulnerability and ecological fragility building bridges between social and natural sciences using the Irish potato famine as a case study. *Conservation Ecology*, 7(2).

Fribourg, C. E. (1977). Andean potato calico strain of tobacco ringspot virus. *Phytopathology*, 67(2):174–178.

Fuchs, M., Hily, J.-M., Petrzik, K., Sanfaçon, H., Thompson, J. R., van der Vlugt, R., Wetzel, T., and Consortium, I. R. (2022). ICTV virus taxonomy profile: Secoviridae 2022. *Journal of General Virology*, 103(12).

Fuchs, M., Schmitt-Keichinger, C., and Sanfaçon, H. (2017). A renaissance in nepovirus research provides new insights into their molecular interface with hosts and vectors. *Advances in Virus Research*, 97:61–105.

Fuentes, S., Gibbs, A. J., Hajizadeh, M., Perez, A., Adams, I. P., Fribourg, C. E., Kreuze, J., Fox, A., Boonham, N., and Jones, R. A. C. (2021). The phylogeography of potato virus X shows the fingerprints of its human vector. *Viruses*, 13(4):644.

Fuentes, S., Heider, B., Tasso, R. C., Romero, E., Zum Felde, T., and Kreuze, J. F. (2012). Complete genome sequence of a potyvirus infecting yam beans (*Pachyrhizus spp.*) in Peru. *Archives of Virology*, 157(4):773–776.

Fuentes, S., Jones, R. A. C., Matsuoka, H., Ohshima, K., Kreuze, J., and Gibbs, A. J. (2019). Potato virus Y; the Andean connection. *Virus Evolution*, 5(2).

Fukuhara, T. (2019). Endornaviruses: persistent dsRNA viruses with symbiotic properties in diverse eukaryotes. *Virus Genes*, 55(2):165–173.

Fukuhara, T., Koga, R., Aoki, N., Yuki, C., Yamamoto, N., Oyama, N., Udagawa, T., Horiuchi, H., Miyazaki, S., Higashi, Y., Takeshita, M., Ikeda, K., Arakawa, M., Matsumoto, N., and Moriyama, H. (2006). The wide distribution of endornaviruses, large double-stranded RNA replicons with plasmid-like properties. *Archives of Virology*, 151(5):995–1002.

Fulton, J. P. (1962). A soil-borne virus in Arkansas. *Arkansas Farm Research*, 11(2):3.

Gaafar, Y. Z., Westenberg, M., Botermans, M., László, K., De Jonghe, K., Foucart, Y., Ferretti, L., Kutnjak, D., Pecman, A., and Mehle, N. (2021). Interlaboratory comparison study on ribodepleted total RNA high-throughput sequencing for plant virus diagnostics and bioinformatic competence. *Pathogens*, 10(9):1174.

Gaafar, Y. Z. A., Herz, K., Hartrick, J., Fletcher, J., Blouin, A. G., MacDiarmid, R., and Ziebell, H. (2020). Investigating the pea virome in Germany—old friends and new players in the field(s). *Frontiers in microbiology*, 11:583242.

Gaafar, Y. Z. A., Richert-Pöggeler, K. R., Sieg-Müller, A., Lüddecke, P., Herz, K., Hartrick, J., Maaß, C., Ulrich, R., and Ziebell, H. (2019). Caraway yellows virus, a novel nepovirus from *Carum carvi*. *Virology Journal*, 16(1).

Gaafar, Y. Z. A. and Ziebell, H. (2020). Comparative study on three viral enrichment approaches based on RNA extraction for plant virus/viroid detection using high-throughput sequencing. *PLoS ONE*, 15(8):e0237951.

Gade, D. W. (1966). Achira, the edible canna, its cultivation and use in the Peruvian Andes. *Economic Botany*, 20(4):407–415.

Gallitelli, D., Mascia, T., and Martelli, G. P. (2012). Viruses in artichoke. *Advances in virus research*, 84:289–324.

Gallitelli, D., Rana, G. L., Vovlas, C., and Martelli, G. P. (2004). Viruses of globe artichoke: an overview. *Journal of Plant Pathology*, 86(4):267–281.

García, M. L., Bó, E. D., da Graça, J. V., Gago-Zachert, S., Hammond, J., Moreno, P., Natsuaki, T., Pallás, V., Navarro, J. A., Reyes, C. A., Luna, G. R., Sasaya, T., Tzanetakis, I. E., Vaira, A. M., Verbeek, M., and Consortium, I. R. (2017). ICTV virus taxonomy profile: Ophioviridae. *Journal of General Virology*, 98(6):1161–1162.

García, M. L., Derrick, K. S., and Grau, O. (1993). Citrus psorosis associated virus and citrus ringspot virus belongs to a new virus group. In *International Organization of Citrus Virologists Conference Proceedings (1957-2010)*, volume 12.

García, W. and Cadima, X. (2003). *Manejo sostenible de la agrobiodiversidad de tubérculos andinos: Síntesis de investigaciones y experiencias en Bolivia*, volume 1. International Potato Center.

García, Y. G., Montoya, M. M., and Gutiérrez, P. A. (2020). Detection of RNA viruses in Cape gooseberry (*Physalis peruviana* L.) by RNAseq using total RNA and dsRNA inputs. *Archives of Phytopathology and Plant Protection*, 53(9-10):395–413.

García-Arenal, F., Fraile, A., and Malpica, J. M. (2001). Variability and genetic structure of plant virus populations. *Annual Review of Phytopathology*, 39:157–186.

García-Díaz, R. F., Valdez-Hernández, E. F., and Ayvar-Serna, S. (2023). Diversity and distribution of Andean tubers: an agrogeographic analysis.

Gardner, A. G., Vaio, M., Guerra, M., and Emshwiller, E. (2012). Diversification of the american bulb-bearing *Oxalis* (*Oxalidaceae*): Dispersal to North America and modification of the tristylous breeding system. *American Journal of Botany*, 99(1):152–164.

Gauthier, M.-E. A., Lelwala, R. V., Elliott, C. E., Windell, C., Fiorito, S., Dinsdale, A., Whattam, M., Pattemore, J., and Barrero, R. A. (2022). Side-by-side comparison of post-entry quarantine and high throughput sequencing methods for virus and viroid diagnosis. *Biology*, 11(2):263.

Gavrilenko, T., Chukhina, I., Antonova, O., Krylova, E., Shipilina, L., Oskina, N., and Kostina, L. (2023). Comparative analysis of the genetic diversity of chilean cultivated potato based on a molecular study of authentic herbarium specimens and present-day gene bank accessions. *Plants*, 12(1):174.

Gayral, P., Noa-Carrazana, J.-C., Lescot, M., Lheureux, F., Lockhart, B. E. L., Matsumoto, T., Piffanelli, P., and Iskra-Caruana, M.-L. (2008). A single banana streak virus integration event in the banana genome as the origin of infectious endogenous pararetrovirus. *Journal of Virology*, 82(13):6697–6710.

Gendall, H., Seminario, J., Sørensen, M., and Theilade, I. (2019). Unearthing the "lost" Andean root crop "mauka" (*Mirabilis expansa* [Ruíz & Pav.] Standl.). *Economic Botany*, 73(4):443–460.

Gentit, P., Delbos, R.-P., Candresse, T., and Dunez, J. (2001). Characterization of a new nepovirus infecting apricot in Southeastern France: Apricot latent ringspot virus. *European Journal of Plant Pathology*, 107(5):485–494.

George, R. A. T. (2011). *Tropical vegetable production*. CABI, Wallingford, Oxfordshire, UK.

Germishuizen, G. and Meyer, N. L. (2003). *Plants of southern Africa: an annotated checklist*, volume 14. National Botanical Institute Pretoria.

Ghosh, D., Chakraborty, S., Kodamana, H., and Chakraborty, S. (2022). Application of machine learning in understanding plant virus pathogenesis: trends and perspectives on emergence, diagnosis, host-virus interplay and management. *Virology Journal*, 19(1).

Gibbs, A. (2020). Binomial nomenclature for virus species: a long view. *Archives of Virology*, 165(12):3079–3083.

Gibbs, A. J., Hajizadeh, M., Ohshima, K., and Jones, R. A. (2020). The potyviruses: An evolutionary synthesis is emerging. *Viruses*, 12(2):132.

Gibbs, M. J., Koga, R., Moriyama, H., Pfeiffer, P., and Fukuhara, T. (2000). Phylogenetic analysis of some large double-stranded RNA replicons from plants suggests they evolved from a defective single-stranded RNA virus. *Journal of General Virology*, 81(1):227–233.

Gilmer, R. M., Uyemoto, J. K., and Kelts, L. J. (1970). A new grapevine disease induced by tobacco ringspot virus. *Phytopathology*, 60:619–27.

Giltrap, N., Eyre, D., and Reed, P. (2009). Internet sales of plants for planting - an increasing trend and threat? *EPPO Bulletin*, 39(2):168–170.

Glyn-Smith, O. (2009). Crap crops of the incas: My on-off-on affair with high altitude andean root crops 1) oca. https://radix4roots.blogspot.com/2009/01/crap-crops-of-incas-my-on-off-on-affair.html. [Online; accessed 12-April-2024].

Goh, C. J. and Hahn, Y. (2021). Analysis of proteolytic processing sites in potyvirus polyproteins revealed differential amino acid preferences of NIa-Pro protease in each of seven cleavage sites. *PLoS One*, 16(1):e0245853.

Goh, R., Xie, X., Lin, Y., Cheng, H., Raja, J. A., and Yeh, S. (2023). Rapid selection of potyviral cross-protection effective mutants from the local lesion host after nitrous acid mutagenesis. *Molecular Plant Pathology*, 24(8):973–988.

Gokalp, K., Digiaro, M., Cigsar, I., Ghanem-Sabanadzovic, N. A., De Stradis, A., Boscia, D., and Martelli, G. P. (2003). Properties of a previously undescribed nepovirus from South-East Anatolia. *Journal of Plant Pathology*, pages 35–41.

Golden, A. M., Franco, J., Jatala, P., and Astogaza, E. (1983). Description of *Thecavermiculatus andinus n.sp.* (*Meloidoderidae*), a round cystoid nematode from the Andes mountains of Peru. *Journal of Nematology*, 15(3):357–63.

Gomes, S. S. V. S. F., Blawid, R., Costa, G. A., Nagata, T., Madeira, N. R., Inoue-Nagata, A. K., Resende, R. O., and Orílio, A. F. (2016). A novel cythorhabdorivirus in arracacha (*Arracacia Xanthorrhiza*). In *Virus Reviews and Research*, volume 2, page 137. XXVII Brazilian Congress of Virology & XI Mercosur Meeting of Virology.

Gonzalez-Franco, A., Langarica, E., Robles-Hernandez, L., Barrios, A., Pérez-Leal, R., Hernandez, A., and Moreno Perez, L. (2014). Detection of virus affecting chilli pepper crop (*Capsicum annuum* L.) in Chihuahua, Mexico. 32:38–51.

Gou, B., Dai, Z., Qin, L., Wang, Y., Liu, H., Wang, L., Liu, P., Ran, M., Fang, C., Zhou, T., Shen, W., Valli, A. A., and Cui, H. (2023). A zinc finger motif in the P1 N terminus, highly conserved in a subset of potyviruses, is associated with the host range and fitness of Telosma mosaic virus. *Journal of Virology*, 97(2):e0144422.

Govaerts, R., Nic Lughadha, E., Black, N., Turner, R., and Paton, A. (2021). The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Scientific Data*, 8(1):215.

Govindarajulu, R., Hughes, C. E., and Bailey, C. D. (2011). Phylogenetic and population genetic analyses of diploid *Leucaena* (*Leguminosae*; *Mimosoideae*) reveal cryptic species diversity and patterns of divergent allopatric speciation. *American Journal of Botany*, 98(12):2049–2063.

Graham, J., Hein, I., and Powell, W. (2007). *Fruits and Nuts*, chapter Raspberry, pages 207–216. Springer.

Granillo, C. R. and Smith, S. H. (1974). Tobacco and tomato ringspot viruses and their relationships with *Tetranychus urticae*. *Phytopathology*, 64(4):494.

Grau, A., Dueñas, R. O., Cabrero, C. N., and Hermann, M. (2003). *Mashua (Tropaeolum tuberosum Ruiz & Pav.) Promoting the conservation and use of underutilized and neglected crops 25*. Promoting the conservation and use of underutilized and neglected crops. International Potato Center (Centro Internacional de la Papa) (CIP), Lima, Peru.

Grau, A. and Halloy, S. (1994). Feast or famine. *Growing Today, June*, 1994:28–33.

Green, S., Cooke, D. E. L., Dunn, M., Barwell, L., Purse, B., Chapman, D. S., Valatin, G., Schlenzig, A., Barbrook, J., Pettitt, T., Price, C., Pérez-Sierra, A., Frederickson-Matika, D., Pritchard, L., Thorpe, P., Cock, P. J. A., Randall, E., Keillor, B., and Marzano, M. (2021). PHYTO-THREATS: Addressing threats to UK forests and woodlands from *Phytophthora*; identifying risks of spread in trade and methods for mitigation. *Forests*, 12(12):1617.

Greener, J. G., Kandathil, S. M., Moffat, L., and Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23(1):40–55.

Griffin, G. D., Huguelet, J. E., and Nelson, J. W. (1963). *Xiphinema americanum* as a vector of necrotic ringspot virus of blueberry. *Plant Disease*, 47(8):703–704.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., and Hofacker, I. L. (2008). The vienna RNA websuite. *Nucleic Acids Research*, 36(suppl_2):W70–W74.

Guimaraes, R. L. and Flores, H. E. (2005). Tropaeolum mosaic potyvirus (TropMV) reduces yield of Andean mashua (*Tropaeolum tuberosum*) accessions. *HortScience*, 40(5):1405–1407.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3):307–321.

Guo, B., Chen, X., Dang, P., Scully, B. T., Liang, X., Holbrook, C. C., Yu, J., and Culbreath, A. K. (2008). Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. *BMC Developmental Biology*, 8:1–16.

Gupta, R., Kwon, S.-Y., and Kim, S. T. (2018). An insight into the tomato spotted wilt virus (TSWV), tomato and thrips interaction. *Plant Biotechnology Reports*, 12:157–163.

Gurung, S., Mahato, S., Thapa, B., and Chhetri, B. (2018). An introduction to ground apple (*Smallanthus sonchifolius*): a review. *International Journal of Genetics*, 10:401.

Gutaker, R. M., Weiß, C. L., Ellis, D., Anglin, N. L., Knapp, S., Luis Fernández-Alonso, J., Prat, S., and Burbano, H. A. (2019). The origins and adaptation of European potatoes reconstructed from historical genomes. *Nature Ecology & Evolution*, 3(7):1093–1101.

Guyader, S. and Ducray, D. G. (2002). Sequence analysis of *Potato leafroll virus* isolates reveals genetic stability, major evolutionary events and differential selection pressure between overlapping reading frame products. *Journal of General Virology*, 83(7):1799–1807.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., Leduc, R. D., Friedman, N., and Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8):1494–1512.

Haegeman, A., Foucart, Y., De Jonghe, K., Goedefroit, T., Al Rwahnih, M., Boonham, N., Candresse, T., Gaafar, Y. Z. A., Hurtado-Gonzales, O. P., Kogej Zwitter, Z., Kutnjak, D., Lamovšek, J., Lefebvre, M., Malapi, M., Mavrič Pleško, I., Önder, S., Reynard, J.-S., Salavert Pamblanco, F., Schumpp, O., Stevens, K., Pal, C., Tamisier, L., Ulubaş Serçe, , Van Duivenbode, I., Waite, D. W., Hu, X., Ziebell, H., and Massart, S. (2023). Looking beyond virus detection in RNA sequencing data: Lessons learned from a community-based effort to detect cellular plant pathogens and pests. *Plants*, 12(11):2139.

Hamilton, R. I., Edwardson, J. R., Francki, R. I. B., Hsu, H. T., Hull, R., Koenig, R., and Milne, R. G. (1981). Guidelines for the identification and characterization of plant viruses. *Journal of General Virology*, 54(2):223–241.

Hanada, K., Fukumoto, F., Kusunoki, M., Kameya-Iwaki, M., Tanaka, Y., and Iwanami, T. (2006). *Cycas necrotic stunt virus* isolated from gladiolus plants in Japan. *Journal of General Plant Pathology*, 72(6):383–386.

Harahagazwe, D., Condori, B., Barreda, C., Bararyenya, A., Byarugaba, A. A., Kude, D. A., Lung'Aho, C., Martinho, C., Mbiri, D., Nasona, B., Ochieng, B., Onditi, J., Randrianaivoarivony, J. M., Tankou, C. M., Worku, A., Schulte-Geldermann, E., Mares, V., Mendiburu, F. D., and Quiroz, R. Q. (2018). How big is the potato (*Solanum tuberosum* L.) yield gap in Sub-Saharan Africa and why? a participatory approach. *Open Agriculture*, 3(1):180–189.

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., Farré, E. M., Veilleux, R. E., and Buell, C. R. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences*, 114(46):E9999–E10008.

Harrison, B. D. (1957a). Soil transmission of beet ringspot virus to peach (*Prunus persica*). *Nature*, 180(4594):1055–1056.

Harrison, B. D. (1957b). Studies on the host range, properties and mode of transmission of beet ringspot virus. *Annals of Applied Biology*, 45(3):462–472.

Harrison, B. D. (1958). Relationship between beet ringspot, potato bouquet and tomato black ring viruses. *Microbiology*, 18(2):450–460.

Harrison, B. D. (1977). Ecology and control of viruses with soil-inhabiting vectors. *Annual Review of Phytopathology*, 15(1):331–360.

Harrison, B. D., Mowat, W. P., and Taylor, C. E. (1961). Transmission of a strain of tomato black ring virus by *Longidorus elongatus* (*Nematoda*). *Virology*, 14(4):480–485.

Hasbun, S. A. (2009). The precautionary principle in the SPS agreement. *ZEuS Zeitschrift für Europarechtliche Studien*, 12(3):455–490.

Hasiów-Jaroszewska, B., Boezen, D., and Zwart, M. P. (2021). Metagenomic studies of viruses in weeds and wild plants: A powerful approach to characterise variable virus communities. *Viruses*, 13(10):1939.

Hasiów-Jaroszewska, B., Budzyńska, D., Borodynko, N., and Pospieszny, H. (2015). Rapid detection of genetically diverse tomato black ring virus isolates using reverse transcription loop-mediated isothermal amplification. *Archives of Virology*, 160(12):3075–3078.

Henderson, R. G. and Wingard, S. A. (1931). Further studies on tobacco ring spot in Virginia. *Journal of Agricultural Research*, 43(3):191–207.

Henning, T. and Von Raab-Straube, E. (2016). *Oxalidaceae*. https://ww2.bgbm.org/europlusmed/ptaxondetail.asp?namecache=oxalis%20articulata&ptreffk=7400000, Euro+Med Plantbase - the information resource for Euro-Mediterranean plant diversity. [Online, accessed 1-February-2024.].

Hensen, A. J. and Stace-smith, R. (1971). Properties of a virus isolated from Golden Elderberry, *Sambucus nigra aurea*. *Phytopathology*, 61(10):1222–1229.

Henz, G. P. (2002). Doenças da mandioquinha-salsa e sua situação atual no Brasil. *Horticultura Brasileira*, 20(2):135–144.

Hermann, M. Heller, J. (1997). *Andean roots and tubers : ahipa, arracacha, maca and yacon. Promoting the conservation and use of underutilized and neglected crops 21*. Promoting the conservation and use of neglected and underutilized crops. International Plant Genetic Resources Institute, Rome, Italy.

Hermann, M. and Bernet, T. (2009). *The transition of maca from neglect to market prominence: lessons for improving use strategies and market chains of minor crops*, volume 1 of *Agricultural Biodiversity and Livelihoods Discussion Papers*. Bioversity International, Rome, Italy. Online.

Hermann, M. and Heller, J. (1997). Arracacha. *Andean roots and tubers: ahipa, arracacha, maca and yacon*, 21:75–172.

Hernándo Bermejo, J. E. and León, J. (1994). *Neglected crops: 1492 from A different perspective*, volume 26 of *Plant Production and Protection Series*. Food & Agriculture Organization of the United Nations, Rome, Italy.

Hibben, C. R. and Bozarth, R. F. (1972). Identification of an ash strain of tobacco ringspot virus. *Phytopathology*, 62:1023–1029.

Hildebrand, E. M. and Kehr, A. E. (1961). Tobacco ringspot virus in sweet potato and its host range in *Nicotiana*. *Phytopathology*, 51(12):833–836.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58:295–300.

Hily, J. M., Poulicard, N., Kubina, J., Reynard, J. S., Spilmont, A. S., Fuchs, M., Lemaire, O., and Vigne, E. (2021). Metagenomic analysis of nepoviruses: diversity, evolution and identification of a genome region in members of subgroup A that appears to be important for host range. *Archives of Virology*, 166(10):2789–2801.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2017). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2):518–522.

Hodgins, K. A., Lai, Z., Oliveira, L. O., Still, D. W., Scascitelli, M., Barker, M. S., Kane, N. C., Dempewolf, H., Kozik, A., Kesseli, R. V., Burke, J. M., Michelmore, R. W., and Rieseberg, L. H. (2014). Genomics of compositae crops: reference transcriptome assemblies and evidence of hybridization with wild relatives. *Molecular Ecology Resources*, 14(1):166–177.

Hoffmann, K., Verbeek, M., Romano, A., Dullemans, A. M., van den Heuvel, J. F. J. M., and van der Wilk, F. (2001). Mechanical transmission of poleroviruses. *Journal of Virological Methods*, 91(2):197–201.

Hollings, M. and Stone, O. M. (1964). Investigation of carnation viruses: I. carnation mottle. *Annals of Applied Biology*, 53(1):103–118.

Horváth, J., Tóbiás, I., and Hunyadi, K. (1994). New natural herbaceous hosts of grapevine fanleaf nepovirus. *Kertészeti Tudomány*, 26(1):31–32.

Hou, W., Li, S., and Massart, S. (2020). Is there a "biological desert" with the discovery of new plant viruses? a retrospective analysis for new fruit tree viruses. *Front Microbiol*, 11:592816.

Hughes, P. L. and Scott, S. W. (2003). First report of *Tomato ringspot virus* in butterfly bush (*Buddleia davidii*). *Plant disease*, 87(1):102–102.

Hull, R. (2001). *Matthews' Plant Virology*. Academic Press, San Diego, CA, 4 edition.

Hulme, P. E. (2014). *The Handbook of Plant Biosecurity*, chapter An Introduction to Plant Biosecurity: Past, Present and Future, pages 1–25. Springer Netherlands.

Humair, F., Humair, L., Kuhn, F., and Kueffer, C. (2015). E-commerce trade in invasive plants. *Conservation Biology*, 29(6):1658–1665.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., and Tappu, R. (2016). MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS computational biology*, 12(6):e1004957.

Icochea, T., Perez, W., and Torres, H. (1995). First report of black rot of *Oxalis tuberosa* tubers caused by *Lasiodiplodia theobromae*. *Plant Disease*, 79(4).

Ilyas, M., Du, Z., and Simon, A. E. (2021). Opium poppy mosaic virus has an Xrn-resistant, translated subgenomic RNA and a BTE 3′ CITE. *Journal of Virology*, 95(9):10.1128/jvi.02109–20.

Ilyas, R., Rohde, M. J., Richert-Pöggeler, K. R., and Ziebell, H. (2022). To be seen or not to be seen: latent infection by tobamoviruses. *Plants*, 11(16):2166.

Inoue-Nagata, A. K., Jordan, R., Kreuze, J., Li, F., López-Moya, J. J., Mäkinen, K., Ohshima, K., Wylie, S. J., and Consortium, I. R. (2022). ICTV virus taxonomy profile: *Potyviridae* 2022. *Journal of General Virology*, 103(5).

Inoue-Nagata, A. K., Oliveira, P. A., Dutra, L. S., and Nagata, T. (2006). Bidens mosaic virus is a member of the *Potato virus Y* species. *Virus Genes*, 33(1):45–49.

International Potato Centre (1999). *CIP Program Report 1997-1998*. International Potato Centre, Lima, Perú.

Invasive Species Advisory Committee (2012). Invasive species and e-Commerce. Report for U.S. Department of the Interior, Washington, DC. https://www.doi.gov/sites/doi.gov/files/uploads/isac_ecommerce_white_paper.rev_.pdf.

IPPC Secretariat (2012). The Internet trade (e-Commerce) in plants: Potential phytosanitary risks. Report, IPPC/FAO, Rome, Italy.

IPPC Secretariat (2013). Pest risk analysis for quarantine pests: International Standard for Phytosanitary Measures No. 11. Standard ISPM No. 11, FAO on behalf of the Secretariat of the International Plant Protection Convention.

IPPC Secretariat (2019). Framework for pest risk analysis: International Standard for Phytosanitary Measures No. 2. Standard, FAO on behalf of the Secretariat of the International Plant Protection Convention.

IPPC Secretariat (2021). Pest risk analysis for regulated non-quarantine pests: International Standard for Phytosanitary Measures No. 21. Standard, FAO on behalf of the Secretariat of the International Plant Protection Convention.

IPPC Secretariat (2023a). *E-Commerce: A guide to managing the pest risk posed by goods ordered online and distributed through postal and courier pathways*. FAO, Rome.

IPPC Secretariat (2023b). Glossary of phytosanitary terms: International Standard for Phytosanitary Measures No. 5. Standard, FAO on behalf of the Secretariat of the International Plant Protection Convention.

Isogai, M., Miyoshi, K., Watanabe, M., and Yoshikawa, N. (2020). Characterization of horizontal transmission of blueberry latent spherical virus by pollen. *Archives of Virology*, 165(12):2807–2815.

Isogai, M., Tatuto, N., Ujiie, C., Watanabe, M., and Yoshikawa, N. (2012). Identification and characterization of blueberry latent spherical virus, a new member of subgroup C in the genus *Nepovirus*. *Archives of Virology*, 157(2):297–303.

Izadpanah, K., Zaki-Aghl, M., and Rowhani, A. (2003a). Non-vitis hosts of grapevine fanleaf virus and their possible epidemiological significance. In *Proceedings of the 14th Meeting of the ICVG*, volume 12. [As cited in Digiaro et al. 2017].

Izadpanah, K., Zaki-Aghl, M., Zhang, Y. P., Daubert, S. D., and Rowhani, A. (2003b). Bermuda grass as a potential reservoir host for *Grapevine fanleaf virus*. *Plant Disease*, 87(10):1179–1182.

Jalkanen, R., Büttner, C., and von Bargen, S. (2007). *Cherry leaf roll virus* abundant on *Betula pubescens* in Finland. *Silva Fennica*, 41:755–762.

Janssen, S. and Giegerich, R. (2015). The RNA shapes studio. *Bioinformatics*, 31(3):423–425.

Jatala, P. (1988). *Improvement of sweet potato (*Ipomaea batatas*) in East Africa, with some references of other tuber and root crops*, chapter Nematodes in tuber and root crops and their management. International Potato Centre.

Jeger, M., Beresford, R., Bock, C., Brown, N., Fox, A., Newton, A., Vicent, A., Xu, X., and Yuen, J. (2021). Global challenges facing plant pathology: multidisciplinary approaches to meet the food security and environmental challenges in the mid-twenty-first century. *CABI Agriculture and Bioscience*, 2(1).

Jeger, M., Bragard, C., Caffier, D., Candresse, T., Chatzivassiliou, E., Dehnen-Schmutz, K., Grégoire, J., Jaques Miret, J. A., Macleod, A., Navajas Navarro, M., Niere, B., Parnell, S., Potting, R., Rafoss, T., Rossi, V., Urek, G., Van Bruggen, A., Van Der Werf, W., West, J., Winter, S., Hart, A., Schans, J., Schrader, G., Suffert, M., Kertész, V., Kozelska, S., Mannino, M. R., Mosbach-Schulz, O., Pautasso, M., Stancanelli, G., Tramontini, S., Vos, S., and Gilioli, G. (2018). Guidance on quantitative pest risk assessment. *EFSA Journal*, 16(8).

Jiang, P., Shao, J., and Nemchinov, L. G. (2019). Identification of the coding-complete genome of cycas necrotic stunt virus in transcriptomic data sets of alfalfa (*Medicago sativa*). *Microbiology Resource Announcements*, 8(38):10.1128/mra. 00981–19.

Jo, Y., Choi, H., Yoon, J.-Y., Choi, S.-K., and Cho, W. K. (2016). *In silico* identification of *Bell pepper endornavirus* from pepper transcriptomes and their phylogenetic and recombination analyses. *Gene*, 575(2, Part 3):712–717.

Jo, Y., Song, M.-K., Choi, H., Park, J.-S., Lee, J.-W., Cho, W. K., and Kim, K.-H. (2018). *In silico* identification of viruses and viroids infecting grapevine cultivar cabernet sauvignon using a grapevine transcriptome. *Journal of Plant Pathology*, 100(1):91–96.

Johns, T. (1981). The anu and the maca. *Journal of Ethnobiology*, 1:208–212.

Jones, A. T., McElroy, F. D., and Brown, D. J. F. (1981). Tests for transmission of cherry leaf roll virus using *Longidorus, Paralongidorus* and *Xiphinema* nematodes. *Annals of Applied Biology*, 99(2):143–150.

Jones, A. T. and McGavin, W. J. (1996). A damaging outbreak of arabis mosaic nepovirus in blackcurrant, the occurrence of other nepoviruses in *Ribes* species, and the demonstration that alfalfa mosaic virus is the cause of interveinal white mosaic in blackcurrant. *Annals of Applied Biology*, 129(1):47–55.

Jones, A. T. and Roberts, I. M. (1977). *Rubus* host range, symptomatology and ultrastructural effects of a British isolate of black raspberry necrosis virus. *Annals of Applied Biology*, 86(3):381–387.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., and Nuka, G. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9):1236–1240.

Jones, R. (1981). Oca strain of arracacha virus b from potato in peru. *Plant Disease*, 65(9):753–754.

Jones, R., Fribourg, C., and Koenig, R. (1983). A previously undescribed nepovirus isolated from potato in peru. *Phytopathology*, 73(2):195–198.

Jones, R. A. (2006). Control of plant virus diseases. *Advances in Virus Research*, 67:205–244.

Jones, R. A. C. (1982). Tests for transmission of four potato viruses through potato true seed. *Annals of Applied Biology*, 100(2):315–320.

Jones, R. A. C. (2009). Plant virus emergence and evolution: Origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, 141(2):113–130.

Jones, R. A. C. (2020). Disease pandemics and major epidemics arising from new encounters between indigenous viruses and introduced crops. *Viruses*, 12(12).

Jones, R. A. C. (2021). Global plant virus disease pandemics and epidemics. *Plants*, 10(2):233.

Jones, R. A. C., Boonham, N., Adams, I. P., and Fox, A. (2021). Historical virus isolate collections: An invaluable resource connecting plant virology's pre-sequencing and post-sequencing eras. *Plant Pathology*, 70:235–248.

Jones, R. A. C., Fox, A., Boonham, N., and Adams, I. P. (2019a). Genomic sequence and host range studies reveal considerable variation within the species *Arracacha virus B*. *Archives of Virology*, 164(11):2849–2852.

Jones, R. A. C. and Kenten, R. H. (1978). Arracacha virus A, a newly recognised virus infecting arracacha (*Arracacia xanthorrhiza; Umbelliferae*) in the Peruvian Andes. *Annals of Applied Biology*, 90(1):85–91.

Jones, R. A. C. and Kenten, R. H. (1981). A strain of arracacha virus B infecting oca (*Oxalis tuberosa: Oxalidaceae*) in the Peruvian Andes. *Journal of Phytopathology*, 100(1):88–95.

Jones, R. A. C. and Naidu, R. A. (2019). Global dimensions of plant virus diseases: Current status and future perspectives. *Annual Review of Virology*, 6(1):387–409.

Jones, S., McGavin, W., and MacFarlane, S. (2019b). The complete sequences of two divergent variants of the rhabdovirus raspberry vein chlorosis virus and the design of improved primers for virus detection. *Virus research*, 265:162–165.

Jongeneel, R. and Herzfeld, T. (2012). *The economics of regulation in agriculture: compliance with public and private standards*, chapter Voluntary certification schemes and legal standards in the EU, pages 97–115. CABI Wallingford UK.

Jordan, R., Mollov, D., Guaragna, M. A., and Lockhart, B. (2019). Detection and first report of beet ringspot virus in ornamental *Oxalis* in the United States. *Plant Disease*, 103(7):1800.

Kahn, S. and Pelgrim, W. (2010). The role of the World Trade Organization and the 'three sisters' (the World Organisation for Animal Health, the International plant Protection Convention and the Codex Alimentarius Commission) in the control of invasive alien species and the preservation of biodiversity. *Revue scientifique et technique (International Office of Epizootics)*, 29(2):411–7.

Kaminski, K., Beckers, F., and Unger, J.-G. (2012). Global internet trade of plants - legality and risks. *EPPO Bulletin*, 42(2):171–175.

Kamp, L., Hartung, J., Mast, B., and Graeff-Hönninger, S. (2019). Tuber yield formation and sugar composition of yacon genotypes grown in Central Europe. *Agronomy*, 9(6):301.

Karapetsi, L., Chatzivassiliou, E. K., Katis, N. I., and Maliogka, V. I. (2021). Artichoke yellow ringspot virus as the causal agent of a new viral disease of lettuce: Epidemiology and molecular variability. *Plant Pathology*, 70(3):594–603.

Kassanis, B. (1957). Effects of changing temperature on plant virus diseases. *Advances in Virus Research*, 4:221–241.

Kavi Sidharthan, V., Diksha, D., Singh, R., Choudhary, S., Naika, M. B. N., and Baranwal, V. K. (2023). Identification of two putative novel deltapartitiviruses and an enamovirus in coriander transcriptomes. *Archives of Microbiology*, 205(10).

Kavi Sidharthan, V., Nagendran, K., and Baranwal, V. K. (2022). Exploration of plant transcriptomes reveals five putative novel poleroviruses and an enamovirus. *Virus Genes*, 58(3):244–253.

Kawchuk, L., Jaag, H. M., Toohey, K., Martin, R., Rohde, W., and Prüfer, D. (2002). In planta agroinfection by Canadian and German *Potato leafroll virus* full-length cDNAs. *Canadian journal of plant pathology*, 24(2):239–243.

Kelleher, M. (2009). *Palgrave Advances in Irish History*, chapter The Irish Famine: History and Representation, pages 84–99. Palgrave Macmillan UK.

Keniknight, G. (1960). *Prunus injucunda* Small and *Acer rubrum* L., natural hosts of peach rosette virus. *Plant Disease Reporter*, 44(3):220.

Kenten, R. H. and Jones, R. A. C. (1979). Arracacha virus B, a second isometric virus infecting arracacha (*Arracacia xanthorrhiza; Umbelliferae*) in the Peruvian Andes. *Annals of Applied Biology*, 93(1):31–36.

Khalili, M., Candresse, T., Koloniuk, I., Safarova, D., Brans, Y., Faure, C., Delmas, M., Massart, S., Aranda, M. A., Caglayan, K., Decroocq, V., Drogoudi, P., Glasa, M., Pantelidis, G., Navratil, M., Latour, F., Spak, J., Pribylova, J., Mihalik, D., Palmisano, F., Saponari, A., Necas, T., Sedlak, J., and Marais, A. (2023). The expanding menagerie of *Prunus*-infecting luteoviruses. *Phytopathology*, 113(2):345–354.

Khankhum, S., Bollich, P., and Valverde, R. A. (2013). First report of *Tobacco ringspot virus* infecting Kudzu (*Pueraria montana*) in Louisiana. *Plant Disease*, 97(4):561–561.

Kim, K.-H. and Hemenway, C. (1997). Mutations that alter a conserved element upstream of the potato virus X triple block and coat protein genes affect subgenomic RNA accumulation. *Virology*, 232(1):187–197.

King, A. M. Q., Lefkowitz, E., Adams, M. J., and Carstens, E. B. (2011). *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, volume 9. Elsevier.

King, S. R. (1987). Four endemic Andean tuber crops: Promising food resources for agricultural diversification. *Mountain Research and Development*, 7(1):43–51.

Kis, S., Salamon, P., Kis, V., and Szittya, G. (2017). Molecular characterization of a beet ringspot nepovirus isolated from *Begonia ricinifolia* in Hungary. *Archives of Virology*, 162(11):3559–3562.

Kleczkowski, A. (1949). The transformation of local lesion counts for statistical analysis. *Annals of Applied Biology*, 36(1):139–152.

Klerks, M. M., Lindner, J. L., Vaškova, D., Špak, J., Thompson, J. R., Jelkmann, W., and Schoen, C. D. (2004). Detection and tentative grouping of *Strawberry crinkle virus* isolates. *European Journal of Plant Pathology*, 110:45–52.

Klingler, J. and Zwicky, P. (1981). Reappearance of nematodes of the genus *Xiphinema* and of virus disease symptoms in vines after soil treatments and fallow. *Mitteilungen Klosterneuburg Rebe und Wein, Obstbau und Fruchteverwertung*, 31(3):89–97.

Knierim, D., Menzel, W., and Winter, S. (2019). Immunocapture of virions with virus-specific antibodies prior to high-throughput sequencing effectively enriches for virus-specific sequences. *PLOS ONE*, 14(5):e0216713.

Kobayashi, K., Sekine, K.-T., and Nishiguchi, M. (2014). Breakdown of plant virus resistance: can we predict and extend the durability of virus resistance? *Journal of General Plant Pathology*, 80(4):327–336.

Koike, A., Murata, T., Matsuda, Y., Masuoka, C., Okamoto, C., and Kabata, K. (2009). Cultivation and ensilage of yacon plants (*Smallanthus sonchifolius* [Poepp. & Endl.] H. robinson) and the function of yacon silage. *Grassland Science*, 55(1):6–10.

Komorowska, B., Ptaszek, M., Jarecka-Boncela, A., and Hasiów-Jaroszewska, B. (2018). First report of arabis mosaic virus in rhubarb in Poland. *Plant Disease*, 102(9):1863–1863.

Komuro, Y. and Iwaki, M. (1968). Tobacco ringspot virus isolated from crotalaria (*Crotalaria spectabilis*). *Japanese Journal of Phytopathology*, 34(1):7–15.

Kondo, T. and Fujita, T. (2012). Complete nucleotide sequence and construction of an infectious clone of Chinese yam necrotic mosaic virus suggest that macluraviruses have the smallest genome among members of the family *Potyviridae*. *Archives of virology*, 157:2299–2307.

Korbie, D. J. and Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, 3(9):1452–1456.

Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., and Simon, R. (2009). Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. *Virology*, 388(1):1–7.

Kreuze, J. F., Souza-Dias, J. A. C., Jeevalatha, A., Figueira, A. R., Valkonen, T. P. J., and Jones, C. A. R. (2020a). *The Potato Crop*, chapter Viral Diseases in Potato, pages 389–430. Springer International Publishing.

Kreuze, J. F., Vaira, A. M., Menzel, W., Candresse, T., Zavriev, S. K., Hammond, J., Hyun Ryu, K., and ICTV Report Consortium, . (2020b). ICTV virus taxonomy profile: *Alphaflexiviridae*. *Journal of General Virology*, 101(7):699–700.

Kundu, J. K., Gadiou, S., Schlesingerova, G., Dziakova, M., and Čermák, V. (2015). Emergence of quarantine *Tobacco ringspot virus* in *Impatiens walleriana* in the Czech Republic. *Plant Protection Science*, 51(3):115–123.

Kusunoki, M., Hanada, K., Iwaki, M., Chang, M. U., Doi, Y., and Yora, K. (1986). Cycas necrotic stunt virus, a new member of nepoviruses found in *Cycas revoluta* host range, purification, serology and and some other properties. 52(2):302–311.

Kutnjak, D., Tamisier, L., Adams, I., Boonham, N., Candresse, T., Chiumenti, M., De Jonghe, K., Kreuze, J. F., Lefebvre, M., Silva, G., Malapi-Wight, M., Margaria, P., Mavrič Pleško, I., Mcgreig, S., Miozzi, L., Remenant, B., Reynard, J.-S., Rollin, J., Rott, M., Schumpp, O., Massart, S., and Haegeman, A. (2021). A primer on the analysis of high-throughput sequencing data for detection of plant viruses. *Microorganisms*, 9(4):841.

Kuzmin, I. V., Novella, I. S., Dietzgen, R. G., Padhi, A., and Rupprecht, C. E. (2009). The rhabdoviruses: biodiversity, phylogenetics, and evolution. *Infection, Genetics and Evolution*, 9(4):541–553.

Lafforgue, G., Lefebvre, M., Michon, T., and Elena, S. F. (2024). How do plant RNA viruses overcome the negative effect of Muller's ratchet despite strong transmission bottlenecks? *bioRxiv*, page 2023.08.01.550272.

Laitenberger, K. (2018). The potential of growing Yacon and other Lost Crops of the Incas and Jerusalem Artichoke as crops and functional foods in Ireland. Report, Nuffield Ireland.

Lana, A. F., Peterson, J. F., Rouselle, G. L., and Vrain, T. C. (1983). Association of tobacco ringspot virus with a union incompatibility of apple. *Journal of Phytopathology*, 106(2):141–148.

Landon, A. J. (2007). Bioprospecting and biopiracy in latin america: The case of maca in perú.

Lane, L. C. (1986). *Methods in Enzymology*, volume 118, chapter Propagation and purification of RNA plant viruses, pages 687–696. Elsevier.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948.

Larsen, R. C., Wyatt, S. D., and Druffel, K. L. (2009). The complete nucleotide sequence and genome organization of red clover vein mosaic virus (genus *Carlavirus*). *Archives of Virology*, 154:891–894.

Latourrette, K., Holste, N. M., and Garcia-Ruiz, H. (2021). Polerovirus genomic variation. *Virus Evolution*, 7(2).

Latvala, S., Susi, P., Lemmetty, A., Cox, S., Jones, A. T., and Lehto, K. (1997). *Ribes* host range and erratic distribution within plants of blackcurrant reversion associated virus provide further evidence for its role as the causal agent of reversion disease. *Annals of Applied Biology*, 131(2):283–295.

Lauber, C., Seifert, M., Bartenschlager, R., and Seitz, S. (2019). Discovery of highly divergent lineages of plant-associated astro-like viruses sheds light on the emergence of potyviruses. *Virus Research*, 260:38–48.

Lauber, C. and Seitz, S. (2022). Opportunities and challenges of data-driven virus discovery. *Biomolecules*, 12:1073.

Laviolette, F. A. and Athow, K. L. (1971). Longevity of Tobacco ringspot virus in soybean seed. *Phytopathology*, 61:755.

Lax, P., Doucet, M. E., Gallardo, C., De L'Argentier, S. M., and Bautista, R. (2008). Presence of soil nematodes in Andean tubers. *Nematropica*, pages 87–94.

Le, T. L. V., Do Truong, L., Vu Dinh, T., and Lebailly, P. (2011). Production and processing of canna vermicelli in Hung Yen province toward sustainable development. In *7th ASAE Conference, Hanoi 2011*. https://ageconsearch.umn.edu/record/290446/?v=pdf.

Le Gall, O., Christian, P., Fauquet, C. M., King, A. M. Q., Knowles, N. J., Nakashima, N., Stanway, G., and Gorbalenya, A. E. (2008). *Picornavirales*, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Archives of Virology*, 153(4):715–727.

Lebas, B., Adams, I., Al Rwahnih, M., Baeyen, S., Bilodeau, J., G., Blouin, A. G., Boonham, N., Candresse, T., Chandelier, A., De Jonghe, K., Fox, A., Gaafar, Y. Z. A., Gentit, P., Haegeman, A., Ho, W., Hurtado-Gonzales, O., Jonkers, W., Kreuze, J., Kutjnak, D., Landa, B., Liu, M., Maclot, F., Malapi-Wight, M., Maree, H. J., Martoni, F., Mehle, N., Minafra, A., Mollov, D., Moreira, A., Nakhla, M., Petter, F., Piper, A. M., Ponchart, J., Rae, R., Remenant, B., Rivera, Y., Rodoni, B., Roenhorst, W., J., Rollin, J., Saldarelli, P., Santala, J., Souza-Richards, R., Spadaro, D., Studholme, D. J., Sultmanis, S., Van Der Vlugt, R., Tamisier, L., Trontin, C., Vazquez-Iglesias, I., Vicente, C. S. L., Vossenberg, B. T. L. H., Wetzel, T., Ziebell, H., and Massart, S. (2022). Facilitating the adoption of high-throughput sequencing technologies as a plant pest diagnostic test in laboratories: A step-by-step description. *EPPO Bulletin*. https://doi.org/10.1111/epp.12863.

Lee, Y.-J., Kwak, H.-R., Lee, Y.-K., Kim, M.-K., Choi, H.-S., and Seo, J.-K. (2015). Complete genome sequence of yacon necrotic mottle virus, a novel putative member of the genus *Badnavirus*. *Archives of Virology*, 160(4):1139–1142.

Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., and Smith, D. B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research*, 46(D1):D708–D717.

Leidi, E. O., Sarmiento, R., and Rodrıguez-Navarro, D. N. (2003). Ahipa (*Pachyrhizus ahipa*[Wedd.] Parodi): an alternative legume crop for sustainable production of starch, oil and protein. *Industrial Crops and Products*, 17(1):27–37.

Lemay, M.-A., Sibbesen, J. A., Torkamaneh, D., Hamel, J., Levesque, R. C., and Belzile, F. (2022). Combined use of oxford nanopore and illumina sequencing yields insights into soybean structural variation biology. *BMC biology*, 20(1):53.

Lemmetty, A., Latvala, S., Jones, A. T., Susi, P., McGavin, W. J., and Lehto, K. (1997). Purification and properties of a new virus from black currant, its affinities with nepoviruses, and its close association with black currant reversion disease. *Phytopathology*, 87(4):404–413.

Leon Velarde, C. U., Quiroz, R. A., Valdivia, R. E., Reinoso, J., and Holle, M. (2008). Envolving from farming systems research into a more holistic rural development approach: Experiences in the Andean region. *Natural Resources Management Division Working Paper.*

Letunic, I. and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1):W293–W296.

Levine, J. M. and D'Antonio, C. M. (2003). Forecasting biological invasions with increasing international trade. *Conservation Biology*, 17(1):322–326.

Levy, L. and Gumpf, D. J. (1991). Studies on the psorosis disease of citrus and preliminary characterization of a flexuous virus associated with the disease. In *International Organization of Citrus Virologists Conference Proceedings (1957-2010)*, volume 11.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing Subgroup, . (2009). The sequence alignment/map format and SAMtools. *bioinformatics*, 25(16):2078–2079.

Liang, X.-Z., Lee, B. T. K., and Wong, S.-M. (2002). Covariation in the capsid protein of Hibiscus chlorotic ringspot virus induced by serial passaging in a host that restricts movement leads to avirulence in its systemic host. *Journal of Virology*, 76(23):12320–12324.

Lim, S., Kwon, S.-Y., Lee, J.-H., Cho, H. S., Kim, H.-S., Park, J. M., Lee, S.-H., and Moon, J. S. (2019). Genomic detection and molecular characterization of two distinct isolates of cycas necrotic stunt virus from *Paeonia suffruticosa* and *Daphne odora*. *Virus Genes*, 55(5):734–737.

Lister, R. M. and Murant, A. F. (1967). Seed-transmission of nematode-borne viruses. *Annals of Applied Biology*, 59(1):49–62.

Liu, Y. Y., Cooper, J. I., Edwards, M. L., and Hellen, C. U. T. (1991). A satellite RNA of arabis mosaic nepovirus and its pathological impact. *Annals of Applied Biology*, 118(3):577–587.

Lizárraga, C. (2000). *Raices andinas: manual de capacitacion*, chapter Virus en raíces andinas. CIP, Lima, Peru.

Lizárraga, C., Chuquillanqui, C., and Jayasinghe, U. (1994). A strain of PBRV (potato black ringspot virus) isolated from arracacha (*Arracacia xanthorrhiza*). *Fitopatología*, 29(2):144–149.

Lizárraga, C., Querci, M., Santa Cruz, M., Bartolini, I., and Salazar, L. F. (2000). Other natural hosts of *Potato virus T*. *Plant Disease*, 84(7):736–738.

Lizárraga, C., Santa Cruz, M., López, G., and Fuentes, S. (2001). *Scientist and Farmer: Partners in Research for the 21st Century. Program Report*, chapter Effect of viruses UMV, UVC, PapMV-U, and PLRV on ulluco production and their control, page 381–389. CIP, Lima, Perú.

Lizárraga, C., Santa Cruz, M., Marca, J. L., and Salazar, L. F. (1999). The importance of the viruses infecting *Ullucus tuberosus* Caldas in Peru. *Fitopatología*, 34(1):22–28.

Lizárraga, C., Santa Cruz, M., and Salazar, L. F. (1996a). First report of potato leafroll luteovirus in ulluco (*Ullucus tuberosus* Caldas). *Plant Disease*, 80(3):344.

Lizárraga, C., Santa Cruz, M., and Salazar, L. F. (1997). *The International Potato Center, Program Report 1995-1996*, chapter Progress in identifying viruses infecting Andean root and tuber crops, pages 156–158. International Potato Centre, Lima, Perú.

Lizárraga, C., Santa Cruz, S., and Jayasinghe, U. (1996b). Detection of an isolate of Andean potato latent tymovirus in ulluco (*Ullucus tuberosus* Caldas). *Plant Disease*, 80:344.

Lockhart, B. E., Menke, J., Dahal, G., and Olszewski, N. E. (2000). Characterization and genomic analysis of tobacco vein clearing virus, a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. *Journal of General Virology*, 81(6):1579–1585.

Lockhart, B. E. L. (2005). Three previously unrecorded viral diseases of *Astilbe, Fuschia*, and *Thermopsis* species in Minnesota. *Plant Disease*, 89(7):775–775.

Lockhart, B. E. L. (2006). Occurrence of *Arabis mosaic virus* in Hostas in the United States. *Plant Disease*, 90(6):834–834.

Lone, S., Harboul, N., and Weltevreden, J. W. J. (2021). 2021 European e-Commerce Report. *Hogeschool van Amsterdam*.

Ma, Y., Marais, A., Lefebvre, M., Theil, S., Svanella-Dumas, L., Faure, C., Candresse, T., and Simon, A. E. (2019). Phytovirome analysis of wild plant populations: Comparison of double-stranded RNA and virion-associated nucleic acid metagenomic approaches. *Journal of Virology*, 94(1):e01462–19.

Ma, Y., Svanella-Dumas, L., Julian, C., Galzi, S., Fernandez, E., Yvon, M., Pirolles, E., Lefebvre, M., Filloux, D., Roumagnac, P., and Candresse, T. (2021). Genome characterization and diversity of trifolium virus 1: identification of a novel legume-infecting capulavirus. *Archives of Virology*, 166(9):2573–2578.

Maachi, A., Donaire, L., Hernando, Y., and Aranda, M. A. (2022). Genetic differentiation and migration fluxes of viruses from melon crops and crop edge weeds. *Journal of Virology*, 96(16):e00421–22.

Maaß, H. I. and Klaas, M. (1995). Infraspecific differentiation of garlic (*Allium sativum* L.) by isozyme and RAPD markers. *Theoretical and Applied Genetics*, 91(1):89–97.

Macconaill, L. E., Burns, R. T., Nag, A., Coleman, H. A., Slevin, M. K., Giorda, K., Light, M., Lai, K., Jarosz, M., Mcneill, M. S., Ducar, M. D., Meyerson, M., and Thorner, A. R. (2018). Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics*, 19(1).

Macdiarmid, R., Rodoni, B., Melcher, U., Ochoa-Corona, F., and Roossinck, M. (2013). Biosecurity implications of new technology and discovery in plant virus research. *PLoS Pathogens*, 9(8):e1003337.

Maclot, F., Candresse, T., Filloux, D., Malmstrom, C. M., Roumagnac, P., Van der Vlugt, R., and Massart, S. (2020). Illuminating an ecological blackbox: Using high throughput sequencing to characterize the plant virome across scales. *Frontiers in Microbiology*, 11(2575).

Maclot, F., Mandujano, M., Nakasato, K., Byrne, J., Paudel, S., Guyer, D., and Malmstrom, C. M. (2023). First report of tobacco ringspot virus infecting pawpaw orchard (*Asimina triloba*) in North America. *Plant Disease*, 107(7).

Madden, L. V., Jeger, M. J., and Van Den Bosch, F. (2000). A theoretical assessment of the effects of vector-virus transmission mechanism on plant virus disease epidemics. *Phytopathology*, 90(6):576–594.

Mahmoodi Safa, J., Nasrolahnejad, S., and Mostafavi Neishaburi, F. S. (2016). Detection and identification of tomato ring spot virus in ornamental plants in North Khorasan province. *Journal of Ornamental Plants*, 6(3):141–150.

Maliogka, V. I., Dovas, C. I., Lesemann, D. E., Winter, S., and Katis, N. I. (2006). Molecular identification, reverse transcription-polymerase chain reaction detection, host reactions, and specific cytopathology of *Artichoke yellow ringspot virus* infecting onion crops. *Phytopathology*, 96(6):622–629.

Manrique, I. (2005). *Yacon syrup: Principles and proccessing*, volume 8. International Potato Center.

Mansouri, F. and Ryšánek, P. (2021). Allexivirus: review and perspectives. *Phytopathologia Mediterranea*, 60(3):389–402.

Maree, H. J., Fox, A., Al Rwahnih, M., Boonham, N., and Candresse, T. (2018). Application of HTS for routine plant virus diagnostics: State of the art and challenges. *Frontiers in Plant Science*, 9:1082.

Maricle, K. L. and Gillock, E. T. (2020). Phylogenetic distribution of an endogenous strain of dahlia mosaic virus in members of *Asteraceae*. *Transactions of the Kansas Academy of Science*, 123(1-2):165–177.

Martelli, G. P. and Taylor, C. E. (1989). Advances in disease vector research.

Marthe, F., Schütze, W., Krüger, H., Scholze, P., Krämer, R., and Ryschka, U. (2003). Maca (*Lepidium meyenii*)-cultivation, resistance and composition of secondary metabolites under European conditions. *Schriften zu Genetischen Ressourcen*, page 290.

Martin, D. P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., and Varsani, A. (2011). Recombination in eukaryotic single stranded DNA viruses. *Viruses*, 3(9):1699–1738.

Martin, R. J., Deo, B., and Halloy, S. (1999). Evaluation of new accessions of Oca or New Zealand yam (*Oxalis tuberosa* Mol.). In *Proceedings of the Twenty-ninth Annual Conference, Agronomy Society of New Zealand*, volume 29, pages 35–39. Agronomy New Zealand.

Martin, R. J., Savage, G. P., Deo, B., Halloy, S. R. P., and Fletcher, P. J. (2005). Development of new oca lines in New Zealand. *Acta Horticulturae*, (670):87–92.

Martín, S., García, M. L., Troisi, A., Rubio, L., Legarreta, G., Grau, O., Alioto, D., Moreno, P., and Guerri, J. (2006). Genetic variation of populations of *Citrus psorosis virus*. *Journal of General Virology*, 87(10):3097–3102.

Marzano, M., Dunn, M., and Green, S. (2021). Perceptions of biosecurity-based accreditation in the plant trade: a UK example. *Forests*, 12(12):1741.

Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M., Reynard, J.-S., Rumbou, A., Saldarelli, P., Škorić, D., Vainio, E. J., Valkonen, J. P. T., Vanderschuren, H., Varveri, C., and Wetzel, T. (2017). A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Frontiers in Microbiology*, 8:45.

Massart, S., Chiumenti, M., De Jonghe, K., Glover, R., Haegeman, A., Koloniuk, I., Komínek, P., Kreuze, J., Kutnjak, D., Lotos, L., Maclot, F., Maliogka, V., Maree, H. J., Olivier, T., Olmos, A., Pooggin, M. M., Reynard, J.-S., Ruiz-García, A. B., Safarova, D., Schneeberger, P. H. H., Sela, N., Turco, S., Vainio, E. J., Varallyay, E., Verdin, E., Westenberg, M., Brostaux, Y., and Candresse, T. (2019). Virus detection by high-throughput sequencing of small RNAs: Large-scale performance testing of sequence analysis strategies. *Phytopathology*, 109(3):488–497.

Massart, S., Olmos, A., Jijakli, H., and Candresse, T. (2014). Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research*, 188:90–6.

Matthews, R. E. F. (1993). *Diagnosis of plant virus diseases*. CRC Press, Boca Raton, FL.

Mayo, M. A. and Ziegler-Graff, V. (1996). *Advances in Virus Research*, volume 46, chapter Molecular Biology of Luteoviruses, pages 413–460. Academic Press.

Mayoa, M. A. and Fritsch, C. (1994). A possible consensus sequence for VPg of viruses in the family comoviridae. *FEBS Letters*, 354(2):129–130.

McDaniel, G. L., Buck, G. J., and Ford, R. E. (1971). Isolation of tobacco ringspot virus from rose. *Phytopathology*, 61:45–49.

Mcgeoch, M. A., Genovesi, P., Bellingham, P. J., Costello, M. J., Mcgrannachan, C., and Sheppard, A. (2016). Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biological Invasions*, 18(2):299–314.

McGuire, J. M. and Douthit, L. B. (1978). Host effect on acquisition and transmission of tobacco ringspot virus by *Xiphinema americanum*. *Phytopathology*, 68(3):457–459.

McLeish, M. J., Fraile, A., and García-Arenal, F. (2019). Evolution of plant–virus interactions: host range and virus emergence. *Current Opinion in Virology*, 34:50–55.

McNeill, W. H. (1999). How the potato changed the world's history. *Social Research*, 66(1):67–83.

Meng, C., Chen, J., Peng, J., and Wong, S.-M. (2006). Host-induced avirulence of *Hibiscus chlorotic ringspot virus* mutants correlates with reduced gene-silencing suppression activity. *Journal of General Virology*, 87(2):451–459.

Messieha, M. (1969). Transmission of tobacco ringspot virus by thrips. *Phytopathology*, 59(7):943–945.

Meyerson, L. A. and Mooney, H. A. (2007). Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment*, 5(4):199–208.

Michi, C., Martinez, E., Cantaluppi, A., Bartolacci, C., Falsini, M., and Lischetti, S. (2023). Plant health campaign strategy. *EFSA Supporting Publications*, 20(5).

Mifsud, J. C. O., Gallagher, R. V., Holmes, E. C., and Geoghegan, J. L. (2022). Transcriptome mining expands knowledge of RNA viruses across the plant kingdom. *Journal of Virology*, 96(24):e00260–22.

Miller, W. A. (1999). *Encyclopedia of Virology*, chapter *Luteovirus* (*Luteoviridae*), pages 901–908.

Miras, M., Sempere, R. N., Kraft, J. J., Miller, W. A., Aranda, M. A., and Truniger, V. (2014). Interfamilial recombination between viruses led to acquisition of a novel translation-enhancing RNA element that allows resistance breaking. *New Phytologist*, 202(1):233–246.

Mishra, R., Shteinberg, M., Shkolnik, D., Anfoka, G., Czosnek, H., and Gorovits, R. (2022). Interplay between abiotic (drought) and biotic (virus) stresses in tomato plants. *Molecular Plant Pathology*, 23(4):475–488.

Misra, M. K., Harries, A., and Dadlani, M. (2023). *Role of Seed Certification in Quality Assurance*, pages 267–297. Springer.

Monsion, B., Froissart, R., Michalakis, Y., and Blanc, S. (2008). Large bottleneck size in *Cauliflower mosaic virus* populations during host plant colonization. *PLoS pathogens*, 4(10):e1000174.

Morales-Briones, D. F., Arias, T., Di Stilio, V. S., and Tank, D. C. (2019). Chloroplast primers for clade-wide phylogenetic studies of *Thalictrum*. *Applications in Plant Sciences*, 7(10):e11294.

Moury, B. and Desbiez, C. (2020). Host range evolution of potyviruses: A global phylogenetic analysis. *Viruses*, 12(1):111.

Muhire, B. M., Varsani, A., and Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PloS one*, 9(9):e108277.

Muimba-Kankolongo, A., Chalwe, A., Sisupo, P., and Kang, M. S. (1997). Distribution, prevalence and outlook for control of cassava mosaic disease in Zambia. *Roots*, 4(1):2–7.

Mundt, J. R. (1973). *Characterization of tobacco ringspot virus isolated from Kansas soybean*. Thesis.

Muyle, A., Bachtrog, D., Marais, G. A. B., and Turner, J. M. A. (2021). Epigenetics drive the evolution of sex chromosomes in animals and plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1826):20200124.

Muñoz, L. and Castillo, R. (1991). *Memorias de la II reunion nacional sobre recursos fitogeneticos*, chapter Pruebas preliminares para la germinación de semilla botánica de oca (*Oxalis tuberosa*) y melloco (*Ullucus tuberosus*), pages 97–101.

Márquez, L. M., Redman, R. S., Rodriguez, R. J., and Roossinck, M. J. (2007). A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. *Science*, 315(5811):513–515.

Nancarrow, N., Aftab, M., Hollaway, G., Rodoni, B., and Trębicki, P. (2022). Symptomless turnip yellows virus infection causes grain yield loss in lentil and field pea: A three-year field study in south-eastern Australia. *Frontiers in Plant Science*, 13.

National Research Council, . (1989). *Lost crops of the Incas: little-known plants of the Andes with promise for worldwide cultivation*. National Academies Press, Washington, DC.

Neeraj, V. and Zaidi, A. A. (2008). *Characterization, diagnosis & management of plant viruses. Volume 4: grain crops & ornamentals*, volume 4 of *Plant Pathogens Series*, chapter Pelargonium: viral diseases and management, pages 245–260. Studium Press LLC, Houston, USA. [As cited in CABI Compendium.

Neilson, R., Handley, L. L., Robinson, D., Scrimgeour, C. M., and Brown, D. J. (1999). Natural abundances of 15N and 13C indicating physiological responses in petunia hybrida to infection by longidorid nematodes and nepoviruses. *Nematology*, 1(3):315–320.

Ng, J. C. K. and Perry, K. L. (2004). Transmission of plant viruses by aphid vectors. *Molecular Plant Pathology*, 5(5):505–511.

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, 32(1):268–274.

Nibert, M. L., Manny, A. R., Debat, H. J., Firth, A. E., Bertini, L., and Caruso, C. (2018). A barnavirus sequence mined from a transcriptome of the Antarctic pearlwort *Colobanthus quitensis*. *Archives of Virology*, 163(7):1921–1926.

Novella, I. S., Clarke, D. K., Quer, J., Duarte, E. A., Lee, C. H., Weaver, S. C., Elena, S. F., Moya, A., Domingo, E., and Holland, J. J. (1995). Extreme fitness differences in mammalian and insect hosts after continuous replication of vesicular stomatitis virus in sandfly cells. *Journal of Virology*, 69(11):6805–6809.

Novák, J. B. and Lanzová, J. (1975). Etiology of two virus diseases of lilac (*Syringa vulgaris* L.) occurring in Czechoslovakia. *Biologia Plantarum*, 17(3):226–227.

Novák, J. B. and Lanzová, J. (1977). Purification and properties of arabis mosaic and tomato bushy stunt viruses isolated from lilac (*Syringa vulgaris* L.). *Biologia Plantarum*, 19(4):264–271.

Nunn, N. and Qian, N. (2010). The Columbian exchange: A history of disease, food, and ideas. *Journal of Economic Perspectives*, 24(2):163–188.

Oberlander, K. C., Dreyer, L. L., and Bellstedt, D. U. (2011). Molecular phylogenetics and origins of southern African *Oxalis*. *TAXON*, 60(6):1667–1677.

Ochoa-Corona, F. M., Tang, J., Lebas, B. S. M., Rubio, L., Gera, A., and Alexander, B. J. R. (2010). Diagnosis of *Broad bean wilt virus 1* and*Verbena latent virus* in *Tropaeolum majus* in New Zealand. *Australasian Plant Pathology*, 39(2):120.

Offei, S. K., Ibidapo, O., Brunt, A. A., and Coutts, R. H. A. (1995). Further evidence for the recognition of Ullucus mild mottle virus as a distinct tobamovirus. *Journal of Phytopathology*, 143(9):543–545.

Ojansivu, I., Ferreira, C. L., and Salminen, S. (2011). Yacon, a new source of prebiotic oligosaccharides with a history of safe use. *Trends in Food Science & Technology*, 22(1):40–46.

Okun, D. O., Khamis, F. M., Muluvi, G. M., Ngeranwa, J. J., Ombura, F. O., Yongo, M. O., and Kenya, E. U. (2015). Distribution of indigenous strains of atoxigenic and toxigenic *Aspergillus flavus* and *Aspergillus parasiticus* in maize and peanuts agro-ecological zones of Kenya. *Agriculture & Food Security*, 4:1–10.

Oliveira, L. M., Orílio, A. F., Inoue-Nagata, A. K., Nagata, T., and Blawid, R. (2017). A novel vitivirus-like sequence found in *Arracacia xanthorrhiza* plants by high throughput sequencing. *Archives of Virology*, 162(7):2141–2144.

Olmos, A., Boonham, N., Candresse, T., Gentit, P., Giovani, B., Kutnjak, D., Liefting, L., Maree, H. J., Minafra, A., Moreira, A., Nakhla, M. K., Petter, F., Ravnikar, M., Rodoni, B., Roenhorst, J. W., Rott, M., Ruiz-García, A. B., Santala, J., Stancanelli, G., van der Vlugt, R., Varveri, C., Westenberg, M., Wetzel, T., Ziebell, H., and Massart, S. (2018). High-throughput sequencing technologies for plant pest diagnosis: challenges and opportunities. 48:219–224.

O'Neil, D., Glowatz, H., and Schlumpberger, M. (2013). Ribosomal RNA depletion for efficient use of RNA-seq capacity. *Current Protocols in Molecular Biology*, 103(1):4.19.1–4.19.8.

Orfanidou, C. G., Lotos, L., Tsiolakis, G., Stefanidis, S. K., Tsialtas, J. T., Katis, N. I., and Maliogka, V. I. (2021). Molecular characterization of poleroviruses isolated from oilseed rape in Greece. *Virus Genes*, 57:289–292.

Orlob, G. B. (1971). History of plant pathology in the middle ages. *Annual Review of Phytopathology*, 9(1):7–20.

Orílio, A. F., Blawid, R., Costa, G. A., Gomes, S. S. V. S. F., Nagata, T., Madeira, N. R., Inoue-Nagata, A. K., and Resende, R. O. (2018). High-throughput sequencing reveals a novel closterovirus in arracacha (*Arracacia xanthorrhiza*). *Archives of Virology*, 163(9):2547–2550.

Orílio, A. F., Dusi, A. N., Madeira, N. R., and Inoue-Nagata, A. K. (2009). Characterization of a member of a new Potyvirus species infecting arracacha in Brazil. *Archives of Virology*, 154(2):181–185.

Orílio, A. F., Dusi, A. N., Madeira, N. R., and Inoue-Nagata, A. K. (2017). First report of *Bidens mosaic virus* in arracacha (*Arracacia xanthorriza*) from Brazil. *Plant Disease*, 101(1):262–262.

Orílio, A. F., Lucinda, N., Dusi, A. N., Nagata, T., and Inoue-Nagata, A. K. (2013). Complete genome sequence of arracacha mottle virus. *Archives of Virology*, 158(1):291–295.

Ouertani, R., Savino, V., Minafra, A., Boscia, D., Castellano, M. A., Martelli, G. P., and Greco, N. (1992). Properties of a previously undescribed grapevine nepovirus from Tunisia. *Archives of Virology*, 126(1):107–117.

Pacot-Hiriart, C., Latvala-Kilby, S., and Lehto, K. (2001). Nucleotide sequence of black currant reversion associated nepovirus RNA1. *Virus Research*, 79(1-2):145–152.

Pagán, I. (2018). The diversity, evolution and epidemiology of plant viruses: A phylogenetic view. *Infection, Genetics and Evolution*, 65:187–199.

Pagán, I., Montes, N., Milgroom, M. G., and García-Arenal, F. (2014). Vertical transmission selects for reduced virulence in a plant virus and for increased resistance in the host. *PLoS Pathogens*, 10(7):e1004293.

Pahalawatta, V., Druffel, K., and Pappu, H. (2008). A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. *Virology*, 376(2):253–257.

Panta, A., Tay, D., Gomez, R., Zea, B., Rojas, E., Simon, R., and Roca, W. (2009). Status and impact of *in vitro* conservation of root and tubers at the International Potato Center (CIP). In *15th Triennial of the Symposium of the International Society for Tropical Root Crops (ISTRC)*, pages 2–6. CIP.

Pappu, H. R., Jones, R. A. C., and Jain, R. K. (2009). Global status of tospovirus epidemics in diverse cropping systems: Successes achieved and challenges ahead. *Virus Research*, 141(2):219–236.

Pasin, F., Menzel, W., and Daròs, J. (2019). Harnessed viruses in the age of metagenomics and synthetic biology: an update on infectious clone assembly and biotechnologies of plant viruses. *Plant Biotechnology Journal*, 17(6):1010–1026.

Pasin, F., Tseng, X.-A., Bedoya, L. C., Heydarnejad, J., Deng, T.-C., García, J. A., and Chen, Y.-R. (2018). Streamlined generation of plant virus infectious clones using the pLX mini binary vectors. *Journal of Virological Methods*, 262:48–55.

Pavliuk, L., Udovychenko, K., Riaba, I., and Bublyk, M. (2021). Detection of sour and sweet cherry viruses in Ukraine.

Pearson, H. G. and Hill, R. L. (2014). Risks for New Zealand's biosecurity from internet trading of plants. In *19th Australasian Weeds Conference, Science, Community and Food Security: The Weed Challenge (Tasmanian Weed Society)*, pages 182–185.

Pecman, A., Kutnjak, D., Gutierrez-Aguirre, I., Adams, I., Fox, A., Boonham, N., and Ravnikar, M. (2017). Next generation sequencing for detection and discovery of plant viruses and viroids: Comparison of two approaches. *Frontiers in Microbiology*, 8:1998.

Penfield, S. (2017). Seed dormancy and germination. *Current Biology*, 27(17):R874–R878.

Pinczinger, D., Von Reth, M., Hanke, M. V., and Flachowsky, H. (2020). SSR fingerprinting of raspberry cultivars traded in germany clearly showed that certainty about the genotype authenticity is a prerequisite for any horticultural experiment. *European Journal of Horticultural Science*, 85(2):79–85.

Pissard, A., Ghislain, M., and Bertin, P. (2006). Genetic diversity of the Andean tuber-bearing species, oca (*Oxalis tuberosa* Mol.), investigated by inter-simple sequence repeats. *Genome*, 49(1):8–16.

Poggi Pollini, C., Babini, A. R., Dall'Ara, M., Mirotti, A., Zucchini, G., and Ratti, C. (2014). First report of cherry leaf roll virus in pomegranate (*Punica granatum* L.) plants. *Journal of Plant Pathology*, 96(4, supplement):60–61.

Poirier, E. Z. and Vignuzzi, M. (2017). Virus population dynamics during infection. *Current opinion in virology*, 23:82–87.

Pollard, C. R. J. and Marzano, M. (2023). On a handshake: business-to-business trust in the biosecurity behaviours of the UK live plant trade. *Biological Invasions*, pages 1–17.

Polák, J. (2007). Viruses of blackthorn and road-bordering trees of plum, myrobalan, sweet and sour cherries in the Czech Republic. *Plant Protection Science*, 43(1):1.

Posnette, A. F. (1969). *Tolerance in Virus Infection in Crop Plants*.

Pospieszny, H., Jonczyk, M., and Borodynko, N. (2003). First record of tomato black ring virus [TBRV] in the natural infection of *Cucumis sativus* in Poland. *Journal of Plant Protection Research*, 43(1).

Powell, C. A., Forer, L. B., Stouffer, R. F., Cummins, J. N., Gonsalves, D., Rosenberger, D. A., Hoffman, J., and Lister, R. M. (1984). Orchard weeds as hosts of tomato ringspot and tobacco ringspot viruses. *Plant Disease*, 68(3):242–244.

Price, W. C. (1936). Specificity of acquired immunity from Tobacco-ring-spot diseases. *Phytopathology*, 26(7):665–675.

Price, W. C. (1940). Comparative host ranges of six plant viruses. *American Journal of Botany*, 27(7):530.

Puffinberger, C. W. and Corbett, M. K. (1985). Euonymus chlorotic ringspot disease caused by tomato ringspot virus. *Phytopathology*, 75(4):423–428.

Pérez-Losada, M., Arenas, M., Galán, J. C., Bracho, M. A., Hillung, J., García-González, N., and González-Candelas, F. (2020). High-throughput sequencing (HTS) for the analysis of viral populations. *Infection, Genetics and Evolution*, 80:104208.

Quemada, H., L′ Hostis, B., Gonsalves, D., Reardon, I. M., Heinrikson, R., Hiebert, E. L., Sieu, L. C., and Slightom, J. L. (1990). The nucleotide sequences of the 3′-terminal regions of papaya ringspot virus strains W and P. *Journal of General Virology*, 71(1):203–210.

Quenouille, J., Vassilakos, N., and Moury, B. (2013). *Potato virus Y*: a major crop pathogen that has provided major insights into the evolution of viral pathogenicity. *Molecular Plant Pathology*, 14(5):439–452.

Rakhshandehroo, F., Zamani Zadeh, H. R., Modarresi, A., and Hajmansoor, S. (2006). Occurrence of *Prunus necrotic ringspot virus* and *Arabis mosaic virus* on rose in Iran. *Plant Disease*, 90(7):975–975.

Ramsdell, D. and Myers, R. (1974). Peach rosette mosaic virus, symptomatology and nematodes associated with grapevine 'degeneration'in michigan. *Phytopathology*, 64(9):1174–1178.

Ramsdell, D. and Myers, R. (1978). Epidemiology of peach rosette mosaic virus in a concord grape vineyard. *Phytopathology*, 68(3):447–450.

Ramsdell, D. and Stace-Smith, R. (1981). Physical and chemical properties of the particles and ribonucleic acid of blueberry leaf mottle virus. *Phytopathology*, 71(4):468–472.

Ramsdell, D. C. and Gillett, J. M. (1981). Peach rosette mosaic virus in highbush blueberry. *Plant Disease*, 65(9):757–758.

Rana, G. L., Gallitelli, D., Kyriakopoulou, P. E., Russo, M., and Martelli, G. P. (1980). Host range and properties of artichoke yellow ringspot virus. *Annals of Applied Biology*, 96(2):177–185.

Ranabhat, N. B., Bruce, M. A., Fellers, J. P., and Shoup Rupp, J. L. (2022). A reproducible methodology for absolute viral quantification and viability determination in mechanical inoculations of wheat streak mosaic virus. *Tropical Plant Pathology*, 47(4):553–561.

Raski, D. J. and Hewitt, W. B. (1963). Plant-parasitic nematodes as vectors of plant viruses.

Read, D. A., Muoma, J., and Thompson, G. D. (2020). Metaviromic analysis reveals coinfection of papaya in western Kenya with a unique strain of Moroccan watermelon mosaic virus and a novel member of the family *Alphaflexiviridae*. *Archives of Virology*, 165(5):1231–1234.

Reddy, P. P. (2015). Achira, *Canna edulis*. *Plant protection in tropical root and tuber crops*, pages 281–291.

Reers, M. (2016). Old crops in a new world: Bringing Andean tuber crops to the Pacific Northwest. Honours college thesis, oregon state university.

Reichel, H., Belalcazar Carvajal, S., Munera, G., Perez, R., and Arevalo, E. (1997). The presence of banana streak virus has been confirmed in plantain (*Musa* AAB Simmonds), sugar cane (*Saccahrum officinarum*) and edible canna (*Canna edulis*) in Colombia. *Infomusa*, 6(1):9–12.

Revers, F., Le Gall, O., Candresse, T., Le Romancer, M., and Dunez, J. (1996). Frequent occurrence of recombinant potyvirus isolates. *Journal of General Virology*, 77(8):1953–1965.

Reyes-Proaño, E., Alvarez-Quinto, R., Delgado-Jiménez, J. A., Cornejo-Franco, J. F., Mollov, D., Bejerman, N., and Quito-Avila, D. F. (2022). Genome characterization and pathogenicity of two new *Hyptis pectinata* viruses transmitted by distinct insect vectors. *Phytopathology*, 112(11):2440–2448.

RHS (2024). *Oxalis* 'irish mist' (v). https://www.rhs.org.uk/plants/287883/oxalis-irish-mist-(v)/details. [Online, accessed 18-April-2024.].

Rivarez, M. P. S., Faure, C., Svanella-Dumas, L., Pecman, A., Tušek-Žnidaric, M., Schönegger, D., De Jonghe, K., Blouin, A., Rasmussen, D. A., and Massart, S. (2023a). Diversity and pathobiology of an ilarvirus unexpectedly detected in diverse plants and global sequencing data. *Phytopathology*, 113(9):1729–1744.

Rivarez, M. P. S., Pecman, A., Bačnik, K., Maksimović, O., Vučurović, A., Seljak, G., Mehle, N., Gutiérrez-Aguirre, I., Ravnikar, M., and Kutnjak, D. (2023b). In-depth study of tomato and weed viromes reveals undiscovered plant virus diversity in an agroecosystem. *Microbiome*, 11(1):1–24.

Rivarez, M. P. S., Vučurović, A., Mehle, N., Ravnikar, M., and Kutnjak, D. (2021). Global advances in tomato virome research: Current status and the impact of high-throughput sequencing. *Frontiers in Microbiology*, 12.

Rivers, T. M. (1937). Viruses and Koch's postulates. *Journal of bacteriology*, 33(1):1–12.

Roberts, J. M. K., Ireland, K. B., Tay, W. T., and Paini, D. (2018). Honey bee-assisted surveillance for early plant virus detection. *Annals of Applied Biology*, 173(3):285–293.

Robinet, C., Kehlenbeck, H., Kriticos, D. J., Baker, R. H. A., Battisti, A., Brunel, S., Dupin, M., Eyre, D., Faccoli, M., Ilieva, Z., Kenis, M., Knight, J., Reynaud, P., Yart, A., and Van Der Werf, W. (2012). A suite of models to support the quantitative assessment of spread in pest risk analysis. *PLoS ONE*, 7(10):e43366.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol*, 29(1):24–6.

Robson, M., Chooi, K. M., Blouin, A. G., Knight, S., and Macdiarmid, R. M. (2022). A national catalogue of viruses associated with indigenous species reveals high-throughput sequencing as a driver of indigenous virus discovery. *Viruses*, 14(11):2477.

Rodoni, B. (2009). The role of plant biosecurity in preventing and controlling emerging plant virus disease epidemics. *Virus Research*, 141(2):150–157.

Roggero, P., Ciuffo, M., Vaira, A. M., Accotto, G. P., Masenga, V., and Milne, R. G. (2000). An Ophiovirus isolated from lettuce with big-vein symptoms. *Archives of virology*, 145:2629–2642.

Rong, W., Rollin, J., Hanafi, M., Roux, N., and Massart, S. (2023). Validation of high-throughput sequencing as virus indexing test for *Musa* germplasm: Performance criteria evaluation and contamination monitoring using an alien control. *PhytoFrontiers*, 3(1):91–102.

Roossinck, M. J. (1997). Mechanisms of plant virus evolution. *Annual Review of Phytopathology*, 35(1):191–209.

Roossinck, M. J. (2005). Symbiosis versus competition in plant virus evolution. *Nature Reviews Microbiology*, 3(12):917–924.

Roossinck, M. J. (2010). Lifestyles of plant viruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1548):1899–1905.

Roossinck, M. J. (2011). The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology*, 9(2):99–108.

Roossinck, M. J. (2015). Plants, viruses and the environment: Ecology and mutualism. *Virology*, 479-480:271–7.

Roossinck, M. J. (2019). Viruses in the phytobiome. *Current Opinion in Virology*, 37:72–76.

Roossinck, M. J. and García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current Opinion in Virology*, 10:56–62.

Roossinck, M. J., Martin, D. P., and Roumagnac, P. (2015). Plant virus metagenomics: Advances in virus discovery. *Phytopathology*, 105(6):716–727.

Roossinck, M. J., Sabanadzovic, S., Okada, R., and Valverde, R. A. (2011). The remarkable evolutionary history of endornaviruses. *Journal of General Virology*, 92(11):2674–2678.

Ross, L. N. and Woodward, J. F. (2016). Koch's postulates: an interventionist perspective. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 59:35–46.

Rott, M. E., Gilchrist, A., Lee, L., and Rochon, D. (1995). Nucleotide sequence of tomato ringspot virus RNA1. *Journal of General Virology*, 76(2):465–473.

Rowhani, A., Daubert, S. D., Uyemoto, J. K., Al Rwahnih, M., and Fuchs, M. (2017). American nepoviruses. *Grapevine Viruses: Molecular Biology, Diagnostics and Management*, pages 109–126.

Roy, B. G., Choi, J., and Fuchs, M. F. (2024). Predictive modeling of proteins encoded by a plant virus sheds a new light on their structure and inherent multifunctionality. *Biomolecules*, 14(1):62.

Rubies-Autonell, C. and Turina, M. (1997). Seed transmission of hibiscus latent ringspot virus (HLRSV). *Plant Disease*, 81(9):1082–1084.

Rubio, L., Galipienso, L., and Ferriol, I. (2020). Detection of plant viruses and disease management: Relevance of genetic diversity and evolution. *Frontiers in Plant Science*, 11.

Rumbou, A., Von Bargen, S., and Büttner, C. (2009). A model system for plant-virus interaction—infectivity and seed transmission of *Cherry leaf roll virus* (CLRV) in *Arabidopsis thaliana*. *European journal of plant pathology*, 124:527–532.

Rush, M. C. and Gooding, G. V. (1970). The occurrence of tobacco ringspot virus strains and tomato ringspot virus in hosts indigenous to North Carolina. *Phytopathology*, 60(12):1756–1760.

Russo, M. (1985). Electron microscopy of grapevine virus infections. *Phytopathologia Mediterranea*, pages 144–147.

Russo, M., Rubino, L., De Stradis, A., and Martelli, G. P. (2009). The complete nucleotide sequence of potato virus T. *Archives of Virology*, 154:321–325.

Rybicki, E. P. (2015). A top ten list for economically important plant viruses. *Archives of Virology*, 160(1):17–20.

Ryckebusch, F., Sauvion, N., Granier, M., and Peterschmitt, M. (2022). *Geminivirus : Detection, Diagnosis and Management*, chapter Spotlight on the recently discovered aphid-transmitted geminiviruses, pages 95–105. Academic Press.

Sabanadzovic, S., Wintermantel, W. M., Valverde, R. A., McCreight, J. D., and Aboughanem-Sabanadzovic, N. (2016). Cucumis melo endornavirus: genome organization, host range and co-divergence with the host. *Virus Research*, 214:49–58.

Salazar, L. F. and Harrison, B. D. (1978). Host range, purification and properties of potato virus t. *Annals of Applied Biology*, 89(2):223–235.

Salem, N., Jewehan, A., Aranda, M. A., and Fox, A. (2023). Tomato brown rugose fruit virus pandemic. *Annual Review of Phytopathology*, 61:137–164.

Samarfard, S., Bejerman, N. E., and Dietzgen, R. G. (2018). Distribution and genetic variability of alfalfa dwarf virus, a cytorhabdovirus associated with alfalfa dwarf disease in Argentina. *Virus Genes*, 54:612–615.

Sammons, B. and Barnett, O. W. (1987). Tobacco ringspot virus from squash grown in South Carolina and transmission of the virus through seed of smooth pigweed. *Plant Disease*, 71(6):530–532.

Samuitienė, M. and Navalinskienė, M. (2001). Nepoviruses and their influence on field floriculture. *Biologija*, 4:43–45.

Samuitienė, M., Navalinskienė, M., and Jackevičienė, E. (2008). Arabis mosaic virus on ornamental plants. *Biologija*, 54(4).

Samuitienė, M., Navalinskienė, M., Kompanets, T., and Bysov, A. (2010). Detection of viruses infecting *Phlox paniculata* in botanical gardens of lithuania and ukraine. *Vytauto Didžiojo universiteto Botanikos sodo raštai*, 14:153–159. [And citations therein.].

Sanches, M. M., De Marchi, B. R., Spadotti, D. M., Nozaki, D. N., Pavan, M. A., and Krause-Sakate, R. (2014). Biological and molecular characterisation of bidens mosaic virus supports its assignment as a member of a distinct species in the genus *Potyvirus*. *Archives of Virology*.

Sanfaçon, H. (2022). Re-examination of nepovirus polyprotein cleavage sites highlights the diverse specificities and evolutionary relationships of nepovirus 3C-like proteases. *Archives of Virology*.

Sanfaçon, H., Wellink, J., Le Gall, O., Karasev, A., Van Der Vlugt, R., and Wetzel, T. (2009). *Secoviridae*: a proposed family of plant viruses within the order *Picornavirale*s that combines the families sequiviridae and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradovirus*. *Archives of Virology*, 154(5):899–907.

Sangketkit, C., Savage, G. P., Martin, R. J., Searle, B. P., and Mason, S. L. (2000). Sensory evaluation of new lines of oca (*Oxalis tuberosa*) grown in New Zealand. *Food Quality and Preference*, 11(3):189–199.

Santa Cruz, M. and Jayasinghe, U. (1994). Aislamiento y caracterización parcial de un virus (AV2) de arracacha (*Arracacia xanthorrhiza*). *Fitopatología*, 32(1):9.

Santillan, F. W., Fribourg, C. E., Adams, I. P., Gibbs, A. J., Boonham, N., Kehoe, M. A., Maina, S., and Jones, R. A. C. (2018). The biology and phylogenetics of potato virus isolates from the Andean Region of South America. *Plant Disease*, 102(5):869–885.

Sastry, K. S. (2013a). *Seed-borne plant virus diseases*, chapter Plant Virus Transmission Through Vegetative Propagules (Asexual Reproduction), pages 285–305. Springer India.

Sastry, K. S. (2013b). *Virus Transmission*, pages 75–83. Springer India.

Sastry, K. S., Mandal, B., Hammond, J., Scott, S. W., and Briddon, R. W. (2019). *Encyclopedia of Plant Viruses and Viroids*, chapter *Pelargonium spp.* (Geranium, Storksbills), pages 1751–1764. Springer India.

Sato, T., Tomioka, K., Nakanishi, T., and Koganezawa, H. (1999). Charcoal rot of yacon (*Smallanthus sonchifolius* (Poepp. et Endl.) H. Robinson), oca (oxalis tuberosa molina) and pearl lupin (Tarwi, *Lupinus mutabilis* Sweet) caused by *Macrophomina phaseolina* (Tassi) Goid. *Bulletin of the Shikoku National Agricultural Experiment Station*, (64):1–8.

Scheidegger, U. C. (1989). *VIII Seminario: Métodos y experiencias de investigación agrícola en campos de agricultores*, chapter Experimentos de semilla de papa en campos de agricultores y su aporte al desarrollo de un programa de semilla, pages 132–143. Quito, Ecuador.

Schellenberger, P., Andret-Link, P., Schmitt-Keichinger, C., Bergdoll, M., Marmonier, A., Vigne, E., Lemaire, O., Fuchs, M., Demangeat, G., and Ritzenthaler, C. (2010). A stretch of 11 amino acids in the βB-βc loop of the coat protein of *Grapevine Fanleaf Virus* is essential for transmission by the nematode *Xiphinema index*. *Journal of Virology*, 84(16):7924–7933.

Schindel, D., Bubela, T., Rosenthal, J., Castle, D., Du Plessis, P., and Bye, R. (2015). The new age of the Nagoya Protocol. *Nature Conservation*, 12:43–56.

Schmelzer, K. (1974). Untersuchungen an Viren der Zier- und Wildgehölze: 8. Mitt.: Neue befunde an Forsythia, Hydrangea und Philadelphus sowie Viren und Virosen an Rhamnus, Centaurea, Galvezia, Cistus, Forestiera, Abeliophyllum, Celastrus, Staphylea und Crambe. *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene. Zweite Naturwissenschaftliche Abteilung: Allgemeine, Landwirtschaftliche und Technische Mikrobiologie*, 129(1):139–168.

Schmelzer, K., Schmidt, H. E., and Schmidt, H. B. (1966). Viruskrankheiten und virusverdächtige Erscheinungen an Forstgehölzen. *Archiv für Forstwesen*, 15:107–20.

Schönegger, D., Marais, A., Babalola, B. M., Faure, C., Lefebvre, M., Svanella-Dumas, L., Brázdová, S., and Candresse, T. (2023a). Carrot populations in France and Spain host a complex virome rich in previously uncharacterized viruses. *PLoS ONE*, 18(8):e0290108.

Schönegger, D., Moubset, O., Margaria, P., Menzel, W., Winter, S., Roumagnac, P., Marais, A., and Candresse, T. (2023b). Benchmarking of virome metagenomic analysis approaches using a large, 60+ members, viral synthetic community. *Journal of Virology*, 97(11):e01300–23.

Secretariat of the Convention on Biological Diversity, . (2011). *Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity: text and annex*. UN.

Seminario, J. (2004). *Raíces Andinas: Contribuciones Al Conocimiento Y a La Capacitación*, chapter Aspectos etnobotánicos y productivos del chago, miso o mauka (*Mirabilis espansa* [Ruíz y Pavón] Standley). CIP, Lima, Peru.

Sequeira, R. and Griffin, R. (2014). *The Handbook of Plant Biosecurity*, chapter The Biosecurity Continuum and Trade: Pre-border Operations, pages 119–148. Springer Netherlands.

Sertkaya, G. (2010). *Tomato ringspot nepovirus* (ToRSV) in wild blackberry (*Rubus fruticosus* L.) in Hatay province of Turkey. *Julius-Kühn-Archiv*, (427):201.

Shaffer, C. M., Michener, D. C., Vlasava, N. B., Chotkowski, H., and Tzanetakis, I. E. (2022). Population genetics of cycas necrotic stunt virus and the development of multiplex RT-PCR diagnostics. *Virus Research*, 309:e198655.

Shang, R., Yang, P., Wang, B., and Zhao, Z. (2018). Transcriptome analysis of maca (*Lepidium meyenii*) root at different developmental stages. *Applications in Plant Sciences*, 6(12):e01206.

Shattuck, V. I. (1992). The biology, epidemiology and control of turnip mosaic virus. *Plant Breeding Reviews*, 14:199–238.

Sherpa, A. R., Hallan, V., Pathak, P., and Zaidi, A. A. (2007). Complete nucleotide sequence analysis of *Cymbidium mosaic virus* Indian isolate: further evidence for natural recombination among potexviruses. *Journal of Biosciences*, 32:663–669.

Shiel, P. J. (1985). Detection of tobacco mosaic and tobacco ringspot viruses in herbaceous and woody plants near virus-infected white ash trees in central New York. *Plant Disease*, page 791.

Shujaei, K., Gibbs, A. J., and Hajizadeh, M. (2020). On the evolution and biogeography of Apple stem grooving capillovirus. *AgriRxiv*, (2020):20203318783.

Sicard, A., Michalakis, Y., Gutiérrez, S., and Blanc, S. (2016). The strange lifestyle of multipartite viruses. *PLoS Pathogens*, 12(11):e1005819.

Sidharthan, V. K. and Baranwal, V. K. (2021). Mining of the water hyssop (*Bacopa monnieri*) transcriptome reveals genome sequences of two putative novel rhabdoviruses and a solendovirus. *Archives of Virology*, 166(7):1985–1990.

Sidharthan, V. K., Kalaivanan, N. S., and Baranwal, V. K. (2021a). Discovery of putative novel viruses in the transcriptomes of endangered plant species native to India and China. *Gene*, 786:145626.

Sidharthan, V. K., Kalaivanan, N. S., and Baranwal, V. K. (2021b). Identification of two putative novel RNA viruses in the transcriptome datasets of small cardamom. *Plant Gene*, 27:100305.

Sidharthan, V. K., Rajeswari, V., and Baranwal, V. K. (2022a). Analysis of public domain plant transcriptomes expands the phylogenetic diversity of the family *Secoviridae*. *Virus Genes*, 58(6):598–604.

Sidharthan, V. K., Rajeswari, V., and Baranwal, V. K. (2023). Broadening the host range and genetic diversity of waikaviruses. *Virology*, 582:106–113.

Sidharthan, V. K., Rajeswari, V., Vanamala, G., and Baranwal, V. K. (2022b). Revisiting the amalgaviral landscapes in plant transcriptomes expands the host range of plant amalgaviruses. *Virology*, 577:65–73.

Silvestre, R., Fuentes, S., Risco, R., Berrocal, A., Adams, I. P., Fox, A., Cuellar, W. J., and Kreuze, J. (2020). Characterization of distinct strains of an aphid-transmitted ilarvirus (Fam. *Bromoviridae*) infecting different hosts from South America. *Virus Research*, 282:197944.

Simberloff, D. (2009). The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, 40:81–102.

Simmonds, P., Adams, M. J., Benkő, M., Breitbart, M., Brister, J. R., Carstens, E. B., Davison, A. J., Delwart, E., Gorbalenya, A. E., Harrach, B., Hull, R., King, A. M., Koonin, E. V., Krupovic, M., Kuhn, J. H., Lefkowitz, E. J., Nibert, M. L., Orton, R., Roossinck, M. J., Sabanadzovic, S., Sullivan, M. B., Suttle, C. A., Tesh, R. B., Van Der Vlugt, R. A., Varsani,

A., and Zerbini, F. M. (2017). Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology*, 15(3):161–168.

Skelton, A., Fowkes, A., Adams, I., Buxton-Kirk, A., Harju, V., Forde, S., Ward, R., Kelly, M., Barber, P., and Fox, A. (2018). First report of *Prunus necrotic ringspot virus* and mulberry cryptic virus 1 in mulberry (*Morus alba*) in the United Kingdom. *New Disease Reports*, 37(23):2044–0588.2018.

Slavíková, L., Ibrahim, E., Alquicer, G., Tomašechová, J., Šoltys, K., Glasa, M., and Kundu, J. K. (2022). Weed hosts represent an important reservoir of turnip yellows virus and a possible source of virus introduction into oilseed rape crop. *Viruses*, 14(11):2511.

Sleat, D. and Palukaitis, P. (1992). A single nucleotide change within a plant virus satellite RNA alters the host specificity of disease induction. *The Plant Journal*, 2(1):43–49.

Smith, A. F. (2012). *Potato: A global history*. Reaktion Books, London, UK.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., Pérez-Emán, J., Burney, C. W., Xie, X., and Harvey, M. G. (2014). The drivers of tropical speciation. *Nature*, 515(7527):406–409.

Smith, K. M. (1931). Composite nature of certain potato viruses of the mosaic group. *Nature*, 127(3210):702–702.

Smith, K. M. and Markham, R. (1944). Two new viruses affecting tobacco and other plants. *Phytopathology*, 34(3):324–329.

Soltani, N., Stevens, K. A., Klaassen, V., Hwang, M.-S., Golino, D. A., and Al Rwahnih, M. (2021). Quality assessment and validation of high-throughput sequencing for grapevine virus diagnostics. *Viruses*, 13(6):1130.

Soria, S., Taboada, C., Rojas, R., Evans, T., Damsteegt, V. D., and Kitto, S. (1996). Performance of virus-free and virus-infected plants of mashua (*Tropaeolum tuberosum*, Ruiz & Pavon). *HortScience*, 31(4):631d–631.

Soria, S. L., Vega, R., Damsteegt, V. D., Mcdaniel, L. L., Kitto, S. L., and Evans, T. A. (1998). Occurrence and partial characterization of a new mechanically transmissible virus in mashua from the Ecuadorian highlands. *Plant Disease*, 82(1):69–73.

Sorrentino, R., Alioto, D., Russo, M., and Rubino, L. (2012). Presence of tobacco ringspot virus in *Aeonium spp. Journal of Plant Pathology*, 94(4).

Sperling, C. R. and King, S. R. (1990). *Advances in new crops*, chapter Andean Tuber Crops: Worldwide Potential, pages 428–435. Timber Press, Portland, OR.

Spetz, C., Taboada, A. M., Darwich, S., Ramsell, J., Salazar, L. F., and Valkonen, J. P. T. (2003). Molecular resolution of a complex of potyviruses infecting solanaceous crops at the centre of origin in Peru. *Journal of General Virology*, 84(9):2565–2578.

Steere, R. L. (1959). *Advances in Virus Research*, volume 6, chapter The purification of plant viruses, pages 1–73. Elsevier.

Steinkellner, H., Himmler, G., Sagl, R., Mattanovich, D., and Katinger, H. (1992). Amino-acid sequence comparison of nepovirus coat proteins. *Virus genes*, 6:197–202.

Stevens, M., McGrann, G., Clark, B., and Authority, H. (2008). *Turnip yellows virus* (syn *Beet western yellows virus*): an emerging threat to European oilseed rape production. *HGCA Research Review*, (69):36.

Stobbe, A. H., Melcher, U., Palmer, M. W., Roossinck, M. J., and Shen, G. (2012). Co-divergence and host-switching in the evolution of tobamoviruses. *Journal of General Virology*, 93(2):408–418.

Stone, W. J., Mink, G. I., and Bergeson, G. B. (1962). A new disease of American spearmint caused by tobacco ringspot virus. *Plant Disease Reporter*, 46:623–624.

Susi, P. (2004). Black currant reversion virus, a mite-transmitted nepovirus. *Molecular Plant Pathology*, 5(3):167–73.

Sõmera, M., Fargette, D., Hébrard, E., Sarmiento, C., and ICTV Report Consortium, . (2021). ICTV virus taxonomy profile: *Solemoviridae* 2021. *Journal of General Virology*, 102(12).

Sørensen, M. (1996). *Yam Bean: Pachyrhizus DC.-Promoting the conservation and use of underutilized and neglected crops. 2*, volume 2. Bioversity International.

Taliansky, M. E. and Robinson, D. J. (2003). Molecular biology of umbraviruses: phantom warriors. *Journal of General Virology*, 84(8):1951–1960.

Tang, J., Lebas, B., Liefting, L., Veerakone, S., Wei, T., and Ward, L. (2016). Opium poppy mosaic virus, a new umbravirus isolated from *Papaver somniferum* in New Zealand. *Archives of Virology*, 161(1):197–201.

Tapio, E. (1985). The appearance of soil-borne viruses in finnish plant nurseries ii. *Agricultural and Food Science*, 57(3):167–181.

Taylor, A. S., Knaus, B. J., Grünwald, N. J., and Burgess, T. (2019). Population genetic structure and cryptic species of *Plasmopara viticola* in Australia. *Phytopathology*, 109(11):1975–1983.

Taylor, C. E. and Thomas, P. R. (1968). The association of *Xiphinema diversicaudatum* (Micoletsky) with strawberry latent ringspot and arabis mosaic viruses in a raspberry plantation. *Annals of Applied Biology*, 62(1):147–157.

Temple, C., Blouin, A. G., De Jonghe, K., Foucart, Y., Botermans, M., Westenberg, M., Schoen, R., Gentit, P., Visage, M., Verdin, E., Wipf-Scheibel, C., Ziebell, H., Gaafar, Y. Z. A., Zia, A., Yan, X.-H., Richert-Pöggeler, K. R., Ulrich, R., Rivarez, M. P. S., Kutnjak, D., Vučurović, A., and Massart, S. (2022). Biological and genetic characterization of Physostegia chlorotic mottle virus in Europe based on host range, location, and time. *Plant Disease*, 106(11):2797–2807.

Temple, C., Blouin, A. G., Tindale, S., Steyer, S., Marechal, K., and Massart, S. (2023). High throughput sequencing technologies complemented by growers' perceptions highlight the impact of tomato virome in diversified vegetable farms and a lack of awareness of emerging virus threats. *Frontiers in Sustainable Food Systems*, 7.

Teycheney, P.-Y., Geering, A. D. W., Dasgupta, I., Hull, R., Kreuze, J. F., Lockhart, B., Muller, E., Olszewski, N., Pappu, H., Pooggin, M. M., Richert-Pöggeler, K. R., Schoelz, J. E., Seal, S., Stavolone, L., Umber, M., and ICTV Report Consortium, . (2020). ICTV virus taxonomy profile: *Caulimoviridae*. *Journal of General Virology*, 101(10):1025–1026.

Thekke-Veetil, T., Polashock, J. J., Marn, M. V., Plesko, I. M., Schilder, A. C., Keller, K. E., Martin, R. R., and Tzanetakis, I. E. (2015). Population structure of blueberry mosaic associated virus: Evidence of reassortment in geographically distinct isolates. *Virus Research*, 201:79–84.

Thiele, G. (1999). Informal potato seed systems in the Andes: Why are they important and what should we do with them? *World Development*, 27(1):83–99.

Thitipraphunkul, K., Uttapap, D., Piyachomkwan, K., and Takeda, Y. (2003). A comparative study of edible canna (*Canna edulis*) starch from different cultivars. Part I. Chemical composition and physicochemical properties. *Carbohydrate Polymers*, 53(3):317–324.

Thomas, W. and Procter, C. (1972). Arabis mosaic virus in *Cyphomandra betaceae* Sendt. *New Zealand Journal of Agricultural Research*, 15(2):395–404.

Thomas, W. and Procter, C. H. (1973). Tobacco ringspot virus in horseradish. *New Zealand Journal of Agricultural Research*, 16(2):233–237.

Thomas-Sharma, S., Andrade-Piedra, J., Carvajal Yepes, M., Hernandez Nopsa, J. F., Jeger, M. J., Jones, R. A. C., Kromann, P., Legg, J. P., Yuen, J., Forbes, G. A., and Garrett, K. A. (2017). A risk assessment framework for seed degeneration: Informing an integrated seed health strategy for vegetatively propagated crops. *Phytopathology*, 107(10):1123–1135.

Thresh, J. M. (1964). Association between black currant reversion virus and its gall mite vector (*Phytoptus ribis* Nal.). *Nature*, 202(4937):1085–1087.

Todesco, M. and De Felippes, F. F. (2016). Why benthamiana went viral. *Trends in Plant Science*, 21(1):4–6.

Tomimura, K., Gibbs, A. J., Jenner, C. E., Walsh, J. A., and Ohshima, K. (2003). The phylogeny of *Turnip mosaic virus*; comparisons of 38 genomic sequences reveal a Eurasian origin and a recent 'emergence'in east asia. *Molecular Ecology*, 12(8):2099–2111.

Torrico, A. K., Salazar, S. M., Kirschbaum, D. S., and Conci, V. C. (2018). Yield losses of asymptomatic strawberry plants infected with *Strawberry mild yellow edge virus*. *European Journal of Plant Pathology*, 150(4):983–990.

Tran, N. T., Teo, A. C., Crew, K. S., Campbell, P. R., Thomas, J. E., and Geering, A. D. W. (2021). Genome sequence and geographic distribution of a new nepovirus infecting *Stenotaphrum secundatum* in Australia. *Virus Research*, 305:198554.

Trontin, C., Agstner, B., Altenbach, D., Anthoine, G., Bagińska, H., Brittain, I., Chabirand, A., Chappé, A. M., Dahlin, P., Dreo, T., Freye-Minks, C., Gianinazzi, C., Harrison, C., Jones, G., Luigi, M., Massart, S., Mehle, N., Mezzalama, M., Mouaziz, H., Petter, F., Ravnikar, M., Raaymakers, T. M., Renvoisé, J. P., Rolland, M., Santos Paiva, M., Seddas, S., Vlugt, R., and Vučurović, A. (2021). VALITEST: Validation of diagnostic tests to support plant health. *EPPO Bulletin*, 51(1):198–206.

Tzanetakis, I. E. and Martin, R. R. (2005). New features in the genus *Ilarvirus* revealed by the nucleotide sequence of *Fragaria chiloensis latent virus*. *Virus Research*, 112(1-2):32–37.

Tzanetakis, I. E., Postman, J. D., and Martin, R. R. (2006). Mint virus X: a novel potexvirus associated with symptoms in 'Variegata' mint. *Archives of Virology*, 151(1):143–153.

Ugent, D., Pozorski, S., and Pozorski, T. (1984). New evidence for ancient cultivation of *Canna edulis* in Peru. *Economic Botany*, 38(4):417–432.

Ulloa, J., Redko, F., Catalán, C., and Muschietti, L. (2021). *Medicinal and Aromatic Plants of the World*, chapter *Smallanthus sonchifolius* (Poepp.) H. Rob, pages 475–491. Springer International Publishing.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, 40(15):e115–e115.

Vaira, A. M., Accotto, G. P., Costantini, A., and Milne, R. G. (2003). The partial sequence of RNA 1 of the ophiovirus *Ranunculus white mottle virus* indicates its relationship to rhabdoviruses and provides candidate primers for an ophiovirus-specific rt-PCR test. *Archives of virology*, 148:1037–1050.

Vaira, A. M. and Milne, R. G. (2008). *Encyclopedia of Virology (Third Edition)*, chapter *Ophiovirus*, pages 447–454. Academic Press, Oxford.

Valentova, L., Rejlova, M., Franova, J., and Cmejla, R. (2022). Symptomless infection by strawberry virus 1 (strV-1) leads to losses in strawberry yields. *Plant Pathology*.

Valentová, K. and Ulrichová, J. (2003). *Smallanthus sonchifolius* and *Lepidium meyenii*-prospective Andean crops for the prevention of chronic diseases. *Biomedical Papers*, 147(2):119–130.

Valli, A., Lopez-Moya, J. J., and García, J. A. (2007). Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family potyviridae. *Journal of General Virology*, 88(3):1016–1028.

Van den Bosch, F., Akudibilah, G., Seal, S., and Jeger, M. (2006a). Host resistance and the evolutionary response of plant viruses. *Journal of Applied Ecology*, 43(3):506–516.

Van den Bosch, F., Jeger, M. J., and Gilligan, C. A. (2006b). Disease control and its selection for damaging plant virus strains in vegetatively propagated staple food crops; a theoretical assessment. *Proceedings of the Royal Society B: Biological Sciences*, 274(1606):11–18.

Van Den Neucker, T. and Scheers, K. (2022). Mislabelling may explain why some prohibited invasive aquatic plants are still being sold in belgium. *Knowledge & Management of Aquatic Ecosystems*, (423):8.

Van Der Vlugt, R. A. and Berendsen, M. (2002). Development of a general potexvirus detection method. *European Journal of Plant Pathology*, 108(4):367–371.

Van der Want, J. P. H., Boerjan, M. L., and Peters, D. (1975). Variability of some plant species from different origins and their suitability for virus work. *Netherlands Journal of Plant Pathology*, 81:205–216.

Van Dijk, P. and Bos, L. (1985). Viral dieback of carrot and other umbelliferae caused by the Anthriscus strain of parsnip yellow fleck virus, and its distinction from carrot motley dwarf. *Netherlands Journal of Plant Pathology*, 91(4):169–187.

Van Dijk, P. and Bos, L. (1989). Survey and biological differentiation of viruses of wild and cultivated umbelliferae in the Netherlands. *Netherlands Journal of Plant Pathology*, 95(2):1–34.

Van Dijk, P., Van Der Meer, F. A., and Piron, P. G. M. (1987). Accessions of Australian *Nicotiana* species suitable as indicator hosts in the diagnosis of plant virus diseases. *Netherlands Journal of Plant Pathology*, 93(2):73–85.

Van Halteren, P. (2000). Diagnostics and National Plant Protection Organizations. *EPPO Bulletin*, 30(3-4):357–359.

Van Regenmortel, M. H. V. (2007). Virus species and virus identification: Past and current controversies. *Infection, Genetics and Evolution*, 7(1):133–144.

Van Regenmortel, M. H. V., Ackermann, H.-W., Calisher, C. H., Dietzgen, R. G., Horzinek, M. C., Keil, G. M., Mahy, B. W. J., Martelli, G. P., Murphy, F. A., Pringle, C., Rima, B. K., Skern, T., Vetten, H.-J., and Weaver, S. C. (2013). Virus species polemics: 14 senior virologists oppose a proposed change to the ICTV definition of virus species. *Archives of Virology*, 158(5):1115–1119.

Vargas, O. M., Madriñán, S., and Simpson, B. (2023). Allopatric speciation is more prevalent than parapatric ecological divergence in a recent high-Andean diversification (*Linochilus: Asteraceae*). *PeerJ*, 11:e15479.

Varsani, A., Roumagnac, P., Fuchs, M., Navas-Castillo, J., Moriones, E., Idris, A., Briddon, R. W., Rivera-Bustamante, R., Murilo Zerbini, F., and Martin, D. P. (2017). *Capulavirus* and *Grablovirus*: two new genera in the family *Geminiviridae*. *Archives of Virology*, 162(6):1819–1831.

Veerakone, S., Liefting, L. W., Lebas, B. S. M., and Ward, L. (2012). First report of *Cherry leaf roll virus* in *Hydrangea macrophylla*. *Plant Disease*, 96(3):463–463.

Velasco, L. and Padilla, C. V. (2021). High-throughput sequencing of small RNAs for the sanitary certification of viruses in grapevine. *Frontiers in Plant Science*, 12.

Verchot, J. (2022). *Potato virus X*: A global potato-infecting virus and type member of the *Potexvirus* genus. *Molecular Plant Pathology*, 23(3):315–320.

Verhoeven, A., Kloth, K. J., Kupczok, A., Oymans, G. H., Damen, J., Rijnsburger, K., Jiang, Z., Deelen, C., Sasidharan, R., Van Zanten, M., and Van Der Vlugt, R. A. A. (2023). Arabidopsis latent virus 1, a comovirus widely spread in *Arabidopsis thaliana* collections. *New Phytologist*, 237(4):1146–1153.

Verhoeven, J. T. J., Ding, Y. M., Lesemann, D.-E., Van der Vlugt, R. A. A., and Roenhorst, J. W. (2000). Virus infections in *Gentiana* species. In *ISHS Acta Horticulturae 568: X International Symposium on Virus Diseases of Ornamental Plants*, pages 61–67.

Verrall, B., Green, K., and Pickering, C. M. (2023). Temporal dynamics in alpine snowpatch plants along a snowmelt gradient explained by functional traits and strategies. *Oecologia*, 201(1):155–171.

Viara, A. and Hammond, J. (2020). *Encyclopedia of Virology*, chapter Ophioviruses (*Aspiviridae*), pages 447–454.

Vietmeyer, N. D. (1991). Lost Crops of the *I*ncas. Magazine article, New Zealand Geographic.

Villamor, D. E. V., Ho, T., Al Rwahnih, M., Martin, R. R., and Tzanetakis, I. E. (2019). High throughput sequencing for plant virus detection and discovery. *Phytopathology*, 109(5):716–725.

Villamor, D. E. V., Keller, K. E., Martin, R. R., and Tzanetakis, I. E. (2022). Comparison of high throughput sequencing to standard protocols for virus detection in berry crops. *Plant Disease*, 106(2):518–525.

Vitali, M. S., Sancho, G., and Katinas, L. (2015). A revision of *Smallanthus* (*Asteraceae, Millerieae*), the "yacón" genus. *Phytotaxa*, 214.

Vivek, A. T., Zahra, S., and Kumar, S. (2020). From current knowledge to best practice: A primer on viral diagnostics using deep sequencing of virus-derived small interfering rnas (vsiRNAs) in infected plants. *Methods*, 183:30–37.

Von Bargen, S., Demiral, R., and Büttner, C. (2015). First detection of *Raspberry ringspot virus* in mosaic diseased hybrid roses in Germany. *New Disease Reports*, 32(18):2044–0588.2015.

Von Bargen, S., Langer, J., Robel, J., Rumbou, A., and Büttner, C. (2012). Complete nucleotide sequence of *Cherry leaf roll virus* (CLRV), a subgroup C nepovirus. *Virus Research*, 163(2):678–683.

Vurro, M., Bonciani, B., and Vannacci, G. (2010). Emerging infectious diseases of crop plants in developing countries: impact on agriculture and socio-economic consequences. *Food Security*, 2(2):113–132.

Walker, P. J., Freitas-Astúa, J., Bejerman, N., Blasdell, K. R., Breyta, R., Dietzgen, R. G., Fooks, A. R., Kondo, H., Kurath, G., and Kuzmin, I. V. (2022). ICTV virus taxonomy profile: *Rhabdoviridae* 2022. *Journal of General Virology*, 103(6):001689.

Walsh, J. A. and Jenner, C. E. (2002). Turnip mosaic virus and the quest for durable resistance. *Molecular Plant Pathology*, 3(5):289–300.

Ward, L. I., Delmiglio, C., Hill, C. F., and Clover, G. R. G. (2009). First report of *Tobacco ringspot virus* on *Sophora microphylla*, a native tree of New Zealand. *Plant Pathology*, 58(4):784–784.

Ward, M. (2016). Action against pest spread—the case for retrospective analysis with a focus on timing. *Food Security*, 8(1):77–81.

Waterworth, H. E. (1971). Physical properties and host ranges of viruses latent in and mechanically transmitted from jasmine. *Phytopathology*, 61:228–230.

Waterworth, H. E. (1975). Purification of arabis mosaic virus isolated from a jasmine plant introduction. *Phytopathology*, 65(8):927–928.

Webber, J. (2010). Pest risk analysis and invasion pathways for plant pathogens. *New Zealand Journal of Forestry Science*, 40(Suppl).

Weber, E. and Gut, D. (2005). A survey of weeds that are increasingly spreading in Europe. *Agronomy for Sustainable Development*, 25(1):109–121.

Weller, S. A., Elphinstone, J. G., Smith, N. C., Boonham, N., and Stead, D. E. (2000). Detection of *Ralstonia solanacearum* strains with a quantitative, multiplex, real-time, fluorogenic PCR (TaqMan) assay. *Applied and Environmental Microbiology*, 66(7):2853–2858.

Were, H. K., Winter, S., and Maiss, E. (2004). Viruses infecting cassava in Kenya. *Plant Disease*, 88(1):17–22.

Whattam, M., Clover, G., Firko, M., and Kalaris, T. (2014). *The Handbook of Plant Biosecurity*, chapter The Biosecurity Continuum and Trade: Border Operations, pages 149–188. Springer Netherlands.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., and Federhen, S. (2007). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, 36(suppl_1):D13–D21.

Wingard, S. A. (1928). Hosts and symptoms of ringspot, a viral disease of plants. *Journal of Agricultural Research*, 37(3):127–153.

Woo, E. N. Y. and Pearson, M. N. (2014). Biological and molecular variation of *Cherry leaf roll virus* isolates from *Malus domestica*, *Ribes rubrum*, *Rubus idaeus*, *Rumex obtusifolius* and *Vaccinium darrowii*. *Plant Pathology*, 63(4):838–845.

Wu, L., Zu, X., Wang, S., and Chen, Y. (2012). Sugarcane mosaic virus–long history but still a threat to industry. *Crop Protection*, 42:74–78.

Wylie, S. J., Coutts, B. A., Jones, M. G. K., and Jones, R. A. C. (2008). Phylogenetic analysis of *Bean yellow mosaic virus* isolates from four continents: relationship between the seven groups found and their hosts and origins. *Plant Disease*, 92(12):1596–1603.

Wylie, S. J., Li, H., Saqib, M., and Jones, M. G. K. (2014). The global trade in fresh produce and the vagility of plant viruses: A case study in garlic. *PLoS ONE*, 9(8):e105044.

Xylogianni, E., Margaria, P., Knierim, D., Sareli, K., Winter, S., and Chatzivassiliou, E. K. (2021). Virus surveys in olive orchards in Greece identify olive virus T, a novel member of the genus Tepovirus. *Pathogens*, 10(5):574.

Yang, L. J., Hidaka, M., Masaki, H., and UozuMI, T. (1998). Detection of potato virus Y P1 protein in infected cells and analysis of its cleavage site. *Bioscience, biotechnology, and biochemistry*, 62(2):380–382.

Yao, J. M., Tainter, F. H., Zimmerman, M. T., and Barnett, O. W. (1994). The occurrence of tomato ringspot virus in *Ponthieva racemosa*. *Plant Disease*, 78(9).

Yasmin, T., Nelson, B. D., Hobbs, H. A., McCoppin, N. K., Lambert, K. N., and Domier, L. L. (2017). Molecular characterization of a new soybean-infecting member of the genus *Nepovirus* identified by high-throughput sequencing. *Archives of Virology*, 162:1089–1092.

Yeomans, A., Plant Health Alliance UK, and Timm, A. (2021). AIPH international plant health survey 2021. Report, AIPH, International Association of Horticultural Producers.

Yeşilçöllü, S., Gümüş, M., and Paylan, I. C. (2011). Ege bölgesi çilek üretim alanlarındaki viral etmenlerin tanılanması üzerinde çalışmalar. *The Journal of Turkish Phytopathology*, 40(1-2-3):13–20.

Yin, D., Qian, J., and Zhu, H. (2019). Frontier development in the midst of ecological civilization: Unravelling the production of maca in Yunnan, China. *Geoforum*, 106:144–154.

Yin, J., Wang, Z., Ma, G., and Liu, W. (2023). Complete chloroplast genome and phylogenetic analysis of *Smallanthus sonchifolius* (*Asteraceae*). *Mitochondrial DNA Part B*, 8(8):916–920.

Yin, Y. Y., Lu, X., Li, T., He, Z., Guo, M., and Ding, M. (2015). Detection of *Turnip mosaic virus* infecting maca in yunnan. *Journal of Plant Pathology*, 97:S67–S77.

Young, D. P. (1958). *Oxalis* in the british isles. *Watsonia*, 4(2):51–69.

Youssef, F., Marais, A., Faure, C., Gentit, P., and Candresse, T. (2011). Strategies to facilitate the development of uncloned or cloned infectious full-length viral cDNAs: *Apple chlorotic leaf spot virus* as a case study. *Virology Journal*, 8(1):488.

Zadoks, J. C. (2001). Plant disease epidemiology in the twentieth century: A picture by means of selected controversies. *Plant Disease*, 85(8):808–816.

Zadoks, J. C. (2008). *On the political economy of plant disease epidemics*, chapter The potato murrain on the European continent and the revolutions of 1848, pages 81–119. Wageningen Academic.

Zamfir, A. D., Babalola, B. M., Fraile, A., McLeish, M. J., and García-Arenal, F. (2023). Tobamoviruses show broad host ranges and little genetic diversity among four habitat types of a heterogeneous ecosystem. *Phytopathology*, 113(9):1697–1707.

Zhang, G. and Sanfaçon, H. (2006). Characterization of membrane association domains within the tomato ringspot nepovirus X2 protein, an endoplasmic reticulum-targeted polytopic membrane protein. *Journal of virology*, 80(21):10847–10857.

Zhang, J., Tian, Y., Yan, L., Zhang, G., Wang, X., Zeng, Y., Zhang, J., Ma, X., Tan, Y., and Long, N. (2016). Genome of plant maca (*Lepidium meyenii*) illuminates genomic basis for high-altitude adaptation in the central Andes. *Molecular Plant*, 9(7):1066–1077.

Zhang, L., Yang, C., Yu, C., Yuan, W., and Ma, Q. (2009). Rapid detection of 3 viruses in *Gerbera jamesonii* Bolus and its virus-free by tissue culture. *Acta Phytophylacica Sinica*, 36(3):239–245.

Zhang, T., Breitbart, M., Lee, W. H., Run, J.-Q., Wei, C. L., Soh, S. W. L., Hibberd, M. L., Liu, E. T., Rohwer, F., and Ruan, Y. (2005). RNA viral community in human feces: Prevalence of plant pathogenic viruses. *PLoS Biology*, 4(1):e3.

Zhang, Y., Pei, X., Zhang, C., Lu, Z., Wang, Z., Jia, S., and Li, W. (2012). *De novo* foliar transcriptome of *Chenopodium amaranticolor* and analysis of its gene expression during virus-induced hypersensitive response.

Zheng, Y.-X., Chen, C.-C., and Jan, F.-J. (2011). First report of *Carnation mottle virus* in *Phalaenopsis* orchids. *Plant Disease*, 95(3):354–354.

Zhu, Y. J., Lim, S. T. S., Schenck, S., Arcinas, A., and Komor, E. (2010). RT-PCR and quantitative real-time RT-PCR detection of *Sugarcane yellow leaf virus* (SCYLV) in symptomatic and asymptomatic plants of Hawaiian sugarcane cultivars and the correlation of SCYLV titre to yield. *European Journal of Plant Pathology*, 127(2):263–273.

Zuļģe, N., Gospodaryk, A., and Moročko-Bičevska, I. (2018). Occurrence and genetic diversity of *Blackcurrant reversion virus* found on various cultivated and wild *Ribes* in Latvia. *Plant Pathology*, 67(1):210–220.

Šneideris, D. and Staniulis, J. (2014). Phylogenetic analysis of Lithuanian tomato black ring virus isolates. *Zemdirbyste*, 101:193–198.

Špak, J., Koloniuk, I., and Tzanetakis, I. E. (2021). Graft-transmissible diseases of *Ribes*–pathogens, impact, and control. *Plant Disease*, 105(2):242–250.

Šubíková, V., Kollerová, E., and Slováková, L. (2002). Occurrence of nepoviruses in small fruits and fruit trees in Slovakia. *Plant Protection Science*, 38:367–369.