Identification and characterization of viruses infecting tobacco (*Nicotiana tabacum* L.) in South Africa

By

Bongeka Sylvia Ndaba

(215049358)

BSc Agriculture (University of KwaZulu-Natal)

Submitted in fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In the

Discipline of Plant Pathology

School of Agriculture, Earth, and Environmental Science

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

Republic of South Africa

December 2021

Dissertation abstract

Tobacco (*Nicotiana tabacum L.*) is one of the most economically important crops which contributes more than R17 billion to the national GDP of South Africa each year. Production, yield and quality of tobacco have been seriously affected by a number of factors including the emerging and recurrent plant viruses. A number of viruses are known to infect tobacco resulting in substantial yield losses. Orthotospoviruses belong to the group of emerging viruses which have shown to spread rapidly over the last decade. These viruses are among the top ten pathogens known to infect a number of crops worldwide. In 2019, large thrips populations accompanied with severe necrotic, ringspot and chlorotic symptoms typical of Orthotospovirus infections were observed in tobacco fields across South Africa (SA). Plant material showing these virus-like symptoms were collected from the three provinces of Limpopo, Northwest and Western Cape for laboratory analysis.

A total of 22 leaf samples of different cultivars exhibiting orthotospovirus-like symptoms were first tested using Enzyme-Linked Immunosorbent Assay with Orthotospovirus specific antibodies, of which 19 of the 22 tested positive. The same samples were further tested with a reverse-transcriptase polymerase reaction (RT-PCR) using 2 sets of primers, firstly with orthotospovirus genericprimers, and then with tomato spotted wilt virus (TSWV) specific primers. RT-PCR results showed that 19 of the 22 samples tested positive for Orthotospoviruses and 11 of the 22 samples tested positive for TSWV. Nineteen positive sample PCR products PCR products of the positive samples were sent for Sanger sequencing. The sequences obtained were subjected to Basic Local Alignment Search Tool (BLAST) and two orthotospoviruses were detected; TSWV and groundnut ringspot virus (GRSV). Eleven of the 19 sequences matched with TSWV and eight matched with GRSV. According to our knowledge, this is the first report of GRSV infecting tobacco in SA.

To determine if the symptoms observed on the tobacco plants were due to mixed virus infections, samples were subjected to Next Generation Sequencing (NGS). For this technique the total RNA was extracted from the frozen symptomatic leaf samples using a Quick-RNA™ Plant Miniprep kit (ZYMO Research, USA). To save costs, all the extracted RNA from the 22 samples was mixed together into one sample and sent to Agricultural Research Council (Biotechnology Platform) for NGS library preparation and sequencing. The NGS data was analysed using the online pipeline software, Genome Detective Virus Tool Version 1.133. The results showed that 11 other plant viruses from different genera, namely; tobacco mosaic virus

(TMV), west African asystasia virus 1 (WAAV1), potato virus Y(PVY), tobacco vein clearing virus (TVCV), tomato leaf curl Uganda virus (ToLCUV), petunia vein clearing virus (PVCV), cucumber mosaic virus (CMV), beet mosaic virus (BMV), beet western yellows virus (BWYV), beet cryptic virus 2 (BCV2), and beet cryptic virus 3(BCV3). These viruses that co-infected tobacco with orthotospoviruses resulting in the severe symptoms observed in the field. NGS analysis also detected the presence of a third orthotospovirus; tomato chlorotic spot (TCSV) in the samples. NGS also was able show the presence of partial and complete sequences for the viruses mentioned above with a coverage between 70-100%. The phylogenetic analysis was done to determine the relationship of each these viruses with other nucleotide sequences of the same species or genus from GENEBANK. Viruses were grouped into clusters according to the sequences of the closest relatives in the genus, and they had a nucleotide identity ranging from 56 to 99 percent to their closest species, which indicates the occurrence of some new isolates

The information presented in this study shows that viruses constitute a significant threat to the economic production of tobacco in SA. Several emerging viruses from different genera that infect tobacco in South Africa were detected using NGS. Most of the viruses detected by NGS are being reported on tobacco in SA for the first time. This study demonstrates the importance and effectiveness of using NGS for plant virus identification without any prior knowledge or based on the symptoms of the virus. Unlike other methods used to identify viruses infecting plants, NGS can detect mixed virus infections. The accurate identification of all viruses infecting tobacco at any given time is crucial for developing effective and sustainable control strategies.

Declaration

I, **Bongeka Sylvia Ndaba**, declare that,

I. The research reported in this dissertation, except where otherwise indicated, and is my original work.

II. This dissertation has not been submitted for any degree or examination at any other university.

III. This dissertation does not contain other persons' data, pictures, graphs, or other information unless specifically acknowledged as being sourced from others.

IV. This dissertation does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, their words have been re-written, but their general information has been referenced. Where their exact words have been used, their writing has been placed inside quotation marks and referenced.

V. Where I have reproduced a publication of which I am an author, co-author, or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

VI. This dissertation does not contain text, graphics or tables copied and pasted from the Internet unless specifically acknowledged. The source is detailed in the dissertation and the References sections.

Signed: …… Date: 01 December 2021

(Bongeka Ndaba, Candidate)

Signed: …… Date: 02/12/2021……

(Professor Augustine Gubba, Supervisor)

Signed: ……. Date: ……03/12/2021……

(Prof Paramu L. Mafongoya, Co-supervisor)

Acknowledgments

- First and foremost, I would like to thank everyone who has supported me throughout this journey.
- To everyone who made sure this project was a success, you guys are incredible!!
- To my supervisor, Professor Augustine Gubba, for pushing me beyond what I believe I can do, I thank you.
- To my co-supervisor, Professor Paramu Mafongoya, for financial assistance and words of encouragement, I appreciate you.
- To the plant pathology technicians Siwe and Khulekani for all technical assistance, going the extra mile to make sure I get everything needed for this project, I thank you.
- To Dr. Jacques Ibaba for assisting with the laboratory techniques and for always answering my unending questions, I appreciate you.
- To my colleagues Nobuhle and Eddie Jnr for being the great lab mates, you are the best.
- To Dr. Vaneson 'Joey' Moodley for sharing your knowledge of virology with us and always motivating us to do our best, and being an inspiration that developed my love for Plant Virology, Thanks to you.
- To my aunt, Sebenzile, for being my pillar of strength, my prayer warrior, my best friend, always encouraging me to study further and constantly reminding me that the sky is the limit, NGIYABONGA.
- To the National Research Foundation (NRF) for funding this project. THANK YOU.

I dedicate this dissertation to my late Mom, Thulisile 'Thandaza' Zwane (1978-2009) and my late Grandmother Nompumelelo Zungu (1954-2020) for always supporting my dreams and unconditional love

Table of Contents

Introduction to Dissertation

Tobacco is an economically important crop in South Africa (SA), contributing over R17 billion/annum (FAO, 2021). According to the recent research on tobacco, it appears that it is more than just a 'cigarette or a smoking pipe filler.' It has been found that tobacco is a powerful biofuel producing three times more energy than maize and soybean (Grisan *et al.,* 2016). However, plant viruses are part of the major production constraints affecting tobacco production. Tobacco viruses reported in SA are potato virus Y (PVY), tobacco mosaic virus (TMV), tomato spotted wilt virus (TSWV), cucumber mosaic virus (CMV), and tobacco necrosis virus (TNV). (Paximadis *et al.,* 1997, Swanepoel and Net, 1995; Rey *et al.,* 1999; Rey *et al.,* 2020).

Several techniques can be used to detect viruses, and these techniques are classified into specific and non-specific methods. Specific methods include; enzyme-linked immunosorbent assay (ELISA), Polymerase chain reaction (PCR)/ reverse transcription (RT-PCR), bioassays, and Sanger sequencing. Non-specific methods include next-generation sequencing (NGS), which can efficiently and accurately detect both DNA and RNA viruses without prior knowledge of the disease pathogen (Villamor *et al.,* 2020). Effective and accurate identification techniques remain the focal point in plant virology and help combat and reduce the negative impact on crop production (Flint *et al.,* 2020).

Problem statement

Recently, severe virus infections were reported in major tobacco-growing regions in South Africa, particularly from the Limpopo, Western Cape, and North West provinces. The tobacco showed virus-like symptoms; severe leaf, stem, and interveinal necrosis, yellowing, wilting, leaf curling, and stunted growth. The symptoms observed in the field were accompanied by large thrips populations. Different thrips species are known to transmit Orthotospoviruses, an important group of viruses with a wide host range causing significant yield losses on a number of crops including tobacco (Whitfield *et al.,* 2005). These typical viral symptoms reduce the yield and quality of the leaves, resulting in significant economic losses. Viral disease outbreaks are complicated by several factors, including rapid evolution, recombination of virus genomes, insect vector migration and population changes, unpredictable expansion of viral host ranges, and migration due to crop trade and

climate changes. These factors make it hard to develop effective long-term disease management strategies (Zaidi *et al.,* 2016).

Significance of the study

In the past two decades, viruses were the problem in SA, but the studies only identified a few of the viruses causing disease epidemics. The previous studies only used specific methods based on what is already seen on the infected crop, e.g., tobacco leaf curl disease (TLCD) was associated with viruses from the *Germiniviridae* family and the presence of whiteflies in the field. However, with virus infections symptoms are not always the best indicator of the virus identity as viruses are known to elicit similar symptoms and mixed virus infections are very common in the field. Unlike the traditional detection techniques, e.g., ELISA, PCR, or biologic assays, NGS technology can identify the viruses without prior knowledge of the insect vector's disease or the transmission mode of the virus. Thistechnology will play a major role in accurately identifying the viruses in mixed infection that can be easily missed because of the absence of the insect vector and symptoms. The proper and accurate identification of viruses will contribute towards the development of proper disease control strategies.

Aim and Objectives

The aim of this study is to identify and characterize viruses infecting tobacco (*Nicotiana* sp.) in South Africa based on what has been seen on the field and also look for mixed infection causing viruses using the recent non-specific technology.

Objectives:

- To identify and characterize isolates of Orthotospoviruses infecting tobacco crops using serological, biological assays, and molecular techniques; RT-PCR, sequencing, and phylogenetic analysis.
- To characterize the complete genome segments of Orthotospoviruses identified
- To identify other co-infecting viruses using NGS

Dissertation structure

The dissertation consists of four chapters. Chapter 1 is a literature review which outlines different plant viruses infecting tobacco in SA, their biology, distribution, genome organization, virus diagnosis, and detection methods, control of viral diseases, and the challenges in developing control strategies will be discussed. Chapter 2 focuses on identifying

and characterizing Orthotospoviruses in SA using ELISA, RT-PCR, sequencing, and phylogenetic analysis; their partial nucleocapsid (N) and non-structural movement protein (NSm) genes. Chapter 3 focuses on the identification of viruses in mixed infections and complete genome characterization of Orthotospoviruses identified in Chapter 2. This dissertation ends with Chapter 4, which sums up the major findings, their implications, and recommendations on what can be done in the future.

References

FAOSTAT <http://www.fao.org/faostat/en/#data/QV accessed 26 September 2021

Flint, S.J., Racaniello, V.R., Rall, G.F., Hatziioannou, T. and Skalka, A.M., 2020. *Principles of virology, Volume 2: pathogenesis and control*. John Wiley & Sons, Hoboken, New Jersey.

Grisan, S., Polizzotto, R., Raiola, P., Cristiani, S., Ventura, F., Di Lucia, F., Zuin, M., Tommasini, S., Morbidelli, R., Damiani, F. and Pupilli, F., 2016. Alternative use of tobacco as a sustainable crop for seed oil, biofuel, and biomass. *Agronomy for Sustainable Development*, *36*(4): 1-8.

Paximadis, M., Dusterwald, G., Duyver, L. and Rey, M.E.C., 1997. Preliminary investigation of the causal agent (s) of a disease-causing leaf curl of tobacco in South Africa. *Plant Pathology*, *46*(2): 282-290.

Paximadis, M., Idris, A.M., Torres-Jerez, I., Villarreal, A., Rey, M.E.C. and Brown, J.K., 1999. Characterization of tobacco geminiviruses in the Old and New World. *Archives of Virology*, *144*(4): 703-717.

Rey, C., 2020. The hunt for plant viruses in South Africa: The early days. *South African Journal of Science*, *116*(11-12): 1-8.

Rey, M.E.C., D'Andrea, E., Calvert-Evers, J., Paximadis, M. and Boccardo, G., 1999. Evidence for a phytoreovirus associated with tobacco exhibiting leaf curl symptoms in South Africa. *Phytopathology*, *89*(4): 303-307.

Swanepoel, A. & Nel, L., 1995. Epidemiological aspects of cucumber mosaic virus on flue-cured tobacco in South Africa. *African Plant Protection*, *1*(1): 13-18.

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.

Whitfield, A.E., Ullman, D.E. and German, T.L., 2005. Tospovirus-thrips interactions. *Annual Review in Phytopathology* 43: 459–489

Zaidi, S.S.E.A., Tashkandi, M., Mansoor, S. and Mahfouz, M.M., 2016. Engineering plant immunity: using CRISPR/Cas9 to generate virus resistance. *Frontiers in Plant Science*, *7*: 1673.

Chapter 1: Literature review

1.1. INTRODUCTION

Tobacco (*Nicotiana tabacum L.)* belongs to the *Solanaceae* family, including potato, tomato, pepper, and eggplant (Christenhusz, 2017). It has been one of the most widely distributed industrial crops worldwide, cultivated in over 150 countries, since it was discovered in America during the late 1400s by Christopher Columbus (Pammel, 1911; Randall, 2012). China, India, Brazil, the United States, Turkey, Zimbabwe, Zambia, and Malawi are the world's largest producers (FAO 2021). *N tabacum* is one of over 80 *Nicotiana* species occurring naturally (Christenhusz, 2017). Tobacco is commonly used as a model plant for fundamental biological processes and in molecular science. It is also used as a model for plant disease susceptibility, which it shares with other *Solanaceae* plants, including potato (*Solanum tuberosum*), tomato (*Solanum lycopersicum*), and pepper (*Capsicum annuum*) (Scholthof, 2004). Recently, research has shown the potential of tobacco as a biofuel. To this end, tobacco can produce up to three times the amount of ethanol per acre as maize (*Zea mays)* and three times the oil per acre as soybean (*Glycine max)* (Grisan *et al.,* 2016)

Tobacco is widely grown in South Africa (SA). About 0.3% of South African agricultural land is used for commercial tobacco cultivation; Virginia tobacco is grown mainly in the provinces of Mpumalanga and Limpopo, with smaller quantities of Oriental tobacco being grown in the provinces of the Western and Eastern Cape (AgriSETA2021). More than 1 000 growers in the country, producing about 34-million kilograms every year on about 24 000ha of land (WHO, 2019). Ninety-six percent (96%) of tobacco growers in SA are smallholder farmers (AgriSETA 2021). The producer price for unmanufactured tobacco is approximately 3675 USD/tonne annually (FAO, 2021). In 2019, a small farmer was able to earn a net income of R5000/ha from tobacco cultivation, which is higher than the returns from other crops such as maize.

Since crop cultivation over ten 10000 years ago, farmers have been plagued by many pests and pathogens, resulting in starvation and social upheaval (Flood, 2010). Dissemination of pests and pathogens occurs through natural and anthropogenic processes facilitated by the increasing interconnectedness of the global food chain (Anderson *et al.,* 2004). Viruses are known to be major limiting factors to crop production worldwide. Because of their perpetuation and the heavy yield losses they cause, viruses are considered among the most economically important

parasites of crops (Ross, 1986). Actual losses due to plant viruses are difficult to assess, but estimates indicate total economic damage as high as several billion US dollars per year.

This review will outline different plant viruses infecting tobacco in SA, their biology, distribution, genome organization, virus diagnosis, and detection methods, control of viral diseases, and the challenges in developing control strategies will be discussed.

1.2. VIRUSES INFECTING TOBACCO IN SOUTH AFRICA

Plant viruses cause significant yield and economic losses (Hančinský *et al.,* 2020). Viruses have been reported in SA, including potato virus Y (PVY), tobacco mosaic virus, tomato spotted wilt virus, cucumber mosaic virus (CMV), tobacco necrosis virus, and tobacco leaf curl virus. (Rey *et al.,* 2020).

1.2.1. Potato virus Y (PVY)

PVY has been the only potyvirus reported to infect tobacco in SA (Paximadis *et al.,* 1997). According to ICTV classification, PVY belongs to the genus *Potyvirus* under the family *Potyviridae*. Potyvirus is one of the largest and economically important plant virus genera with more than 190 species, making it the largest genera of plant infecting RNA viruses (ICTV, 2021). PVY has a non-enveloped flexuous rod shape, with the size of between 680-900 nanometers (nm) in length and 12-15 nm in diameter (Fig. 1.1a). It is has a positive sense of a single-stranded linear RNA genome of about 8-11kb in size. The genome consists of a single polyprotein encodes for about ten proteins (Fig. 1.1b; Vijayapalani *et al.,* 2012).

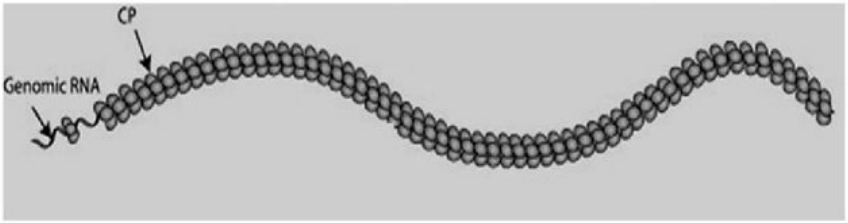


Figure1.1a: Potyvirus flexuous rod particle, with a linear genomic RNA inside the coat protein (CP) (Mishra *et al.,* 2014)

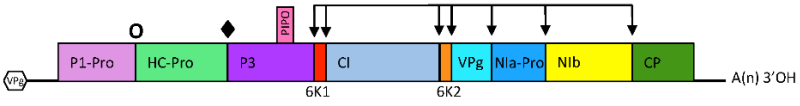


Figure 1.1b: Genomic map of viruses in the genus Potyvirus, with a 5' terminus with a genome-linked protein (VPg) (Wylie *et al.,* 2017).

PVY is an aphid transmitted virus that was first reported in 1931, and it is in the top ten economically important viruses worldwide (Rybicki, 2015; Bellstedt *et al.,* 2017.). It has a wide host range; it infects economically important crops such as potato, pepper, tomato, and tobacco. PVY infects 70 *Nicotiana* species (Bellstedt *et al.,* 2017). PVY consists of different strains classified into three main groups: N, C, and O. Tobacco is commonly infected by PVYO and PVYC, associated with mosaic and chlorotic ring symptoms (Fig. 1.2.b&d) (Guo *et al.,* 2017). PVY^{N}, $PVY^{NTN,}$ and $PVY^{N}W$ strains are common on tobacco, causing veinal necrosis symptoms (Fig. 1.2b) (Glais *et al.,* 2017).

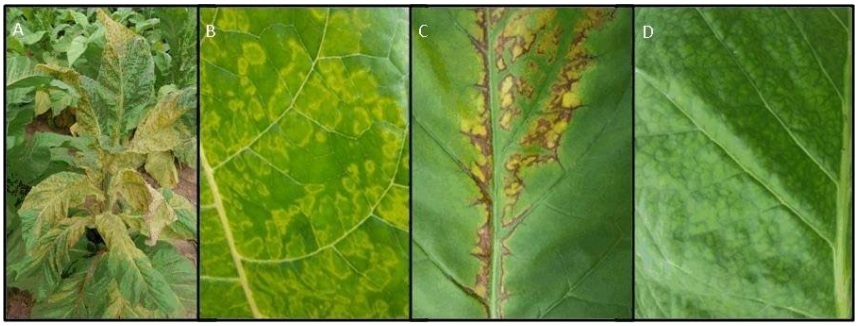


Figure 1.2. : A: Potato Virus Y diseased plant B: Characteristic symptoms showing chlorotic rings C: necrotic veins D: vein banding) (source: The Tobacco Research Board, (TRB)

1.2.2. Tomato spotted wilt virus (TSWV)

TSWV belongs to the *Tospoviridae* family under the genus orthotospovirus, one of the fastest spreading group of plant viruses, causing severe crop damage over the past decade. There are over 29 orthotospovirus species reported up to date (Turina *et al.,* 2016). The first orthotospovirus was first described by Samuel *et al*. (1930) and reviewed by Best (1968). *Tospoviridae* derives its name from the tomato spotted wilt virus, first reported in Victoria, Australia, in 1915 (Brittlebank, 1919). TSWV is among the most detrimental of all known plant viruses. It is one of the top ten economically important viruses (Rybicki, 2015), causing crop yield losses exceeding 1 billion US dollars (Pappu et al., 2009). TSWV has a wide host range infecting over 1000 plant species worldwide.

TSWV particles are quasi-spherical with a diameter of approximately 80-120 nm. Each viral particle consists of a granular core of nucleocapsids, bound by a lipid envelope (5-10 nm thick) which is covered with surface projections that comprised of two glycoproteins, Gn (78 KDa) and Gc (58 KDa) (Fig. 1.3; Pavithra *et al.,* 2016). The virus possesses a linear tripartite negative single-stranded RNA genome of which one segment is of negative polarity, and the other two are ambisense (Kormelink *et al.,* 1992). These three RNA segments are named small (S),

Medium (M), and Large (L) based on their length (Fig.1.4). The TSWV genome codes for six proteins; RNA-dependent RNA polymerase, two structural glycoproteins (Gn and Gc), nonstructural movement protein (Nsm), nucleocapsid protein (N), and nonstructural protein (NSs). (Whitfield et al., 2005; Pappu et al., 2009).

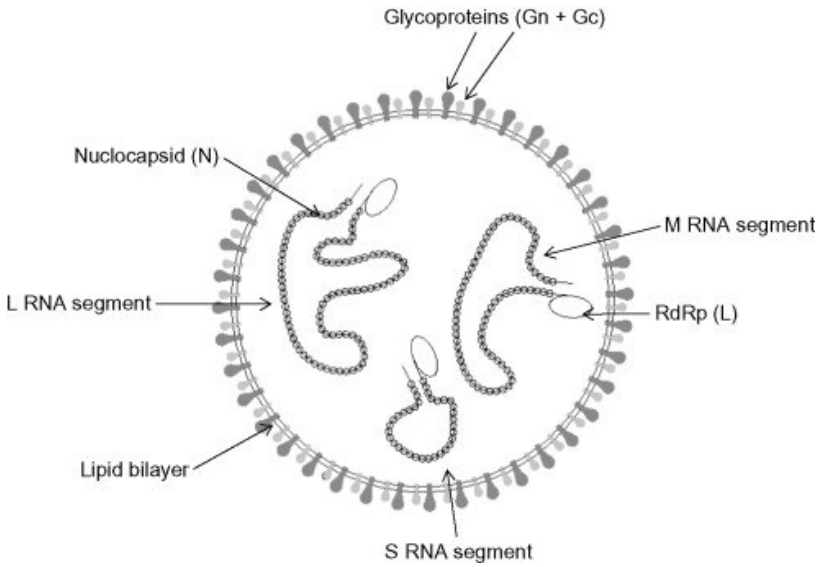


Figure 1.3. Representation of the TSWV virion. The virion comprises a lipid envelope, which encapsulates the three viral genomic RNAs, named according to their small, medium, and large size. The genomic RNA is encapsidated by nucleocapsid (N) protein, which has been shown to interact with the cytoplasmic tails of the membrane-spanning viral glycoproteins (Turina et al., 2012)

TSWV-like symptoms were observed as earlier as 1905 in South Africa (SA)." and it was known as the "wilt disease of tobacco," which was later named as 'Kromnek' (Moore, 1932). It was later shown to be induced by TSWV, and it was the first and only orthotospovirus to infect tobacco in SA. TSWV is transmitted by *Frankliniella schultzei* and *Thrips tabaci* (Thompson and van Zijl, 1996). By 1939, it was reported in several SA provinces (Western Cape, Free State, and throughout the former Transvaal) (Moore and Andessen, 1939), and the spread was reported just after the introduction of the *Frankliniella occidentalis*, the western flower thrips (Wijkamp *et al.,* 1996). TSWV has become a major concern because it infects many economically important crops and weeds that grow along with those crops and act as a green bridge to spread the virus (Moore and Andessen, 1939). The emergence of TSWV is a severe threat to cultivating many essential crops in SA, such as tomatoes, potatoes, pepper, and tobacco (Sivparasad and Gubba, 2008). TSWV causes severe leaf and stems necrosis, chlorosis, and wilting symptoms on tobacco (Fig. 1.5).

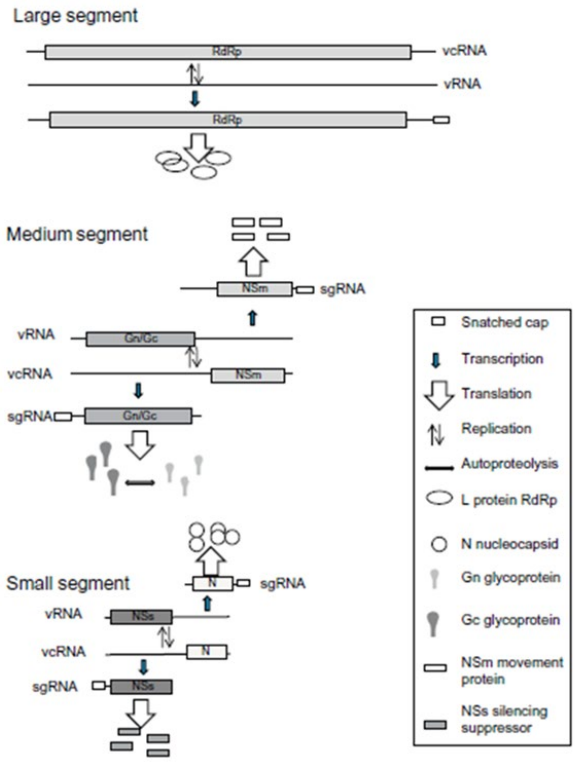


Figure 1.4.: Schematic representation of a tospovirus genome with the main steps for protein synthesis, subgenomic (sg) transcription, and genome replication cycle. V, encapsidated genomic segment; vc, complementary strand to virus encapsidated genomic strand (Turina *et al.,* 2012)

Figure 1.5.: Typical TSWV symptoms on tobacco (A&B); Severe chlorotic lesions on stems leaves, wilting, and plant yellowing (Source: Paul Bachi, University of Kentucky Research and Education Center, Bugwood.org)

1.2.3. Cucumber mosaic virus (CMV)

CMV is one of the top ten economically important plant viruses worldwide, with a wide host range infecting over 1000 plant species in the 85 families ((Liu *et al.,* 2019). CMV belongs to the genus *Cucumovirus* under the family *Bromoviridae.* CMV was first reported in 1916 as a cucumber disease in New York; since then, it has spread worldwide (Korbecka-Glinka et al., 2021). CMV has non-enveloped icosahedral-shaped particles with a size of 29nm in diameter. It has a positive sense single-stranded tripartite RNA genome consisting of RNA1, RNA2, and RNA3 (Canto *et al.,* 1997). The genomic RNA acts as mRNA. RNA1 and RNA2 code for 1a and 2a proteins, respectively that are involved in the replication. Subgenomic RNA4 (sgRNA4) is transcribed from the negative strand copy of RNA3; both sgRNA4 and RNA3 are translated into capsid and movement proteins, respectively (Fig. 1.6; Tungadi et al., 2020). Recombination and reassortment can occur on the CMV genome because of high genetic diversity, although not that frequent (Qiu *et al.,* 2017).

CMV is the only cucumovirus that has been reported to infect tobacco in SA. The highest CMV epidemic on tobacco was reported in the 1987/8 growing season (Swanepoel and Nel, 1995). Over 60 species of aphids transmit CMV, and it is common in areas where vegetables are produced, mostly *Cucurbitaceae* and *Solanaceae* crops (Liu *et al.,* 2019). Typical symptoms of CMV on tobacco include mottling or mosaic on the young leaves, sometimes accompanied by stunted growth, narrowing that often lead to shoestring symptoms, or deformed older leaves(Fig. 1.7; Garcıa-Arenal and Palukaitis 2009). Scalding of infected leaves may also occur. So far, there are no tobacco cultivars resistant to CMV (Nemes *et al.,* 2019).

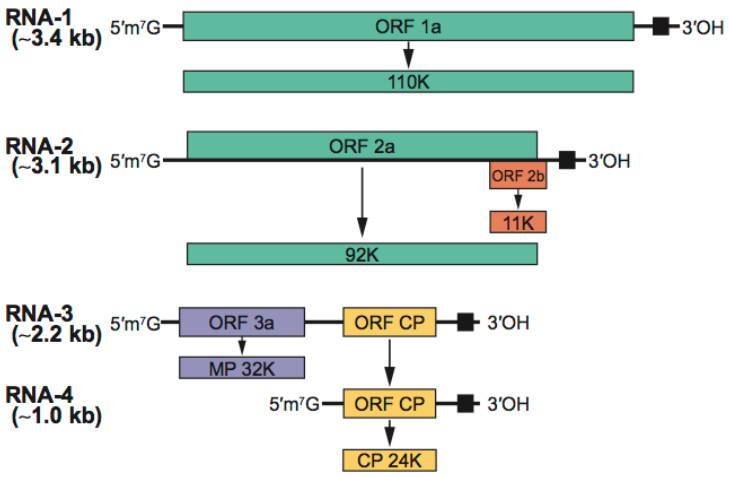


Figure 1.6.: CMV genome segmented, linear tripartite ssRNA (+) composed of RNA1, RNA2, RNA3. Each genomic segment has a 3' tRNA-like structure and a 5'cap. (ICTV, 2021)

Figure 1.7.: CMV typical symptoms on tobacco, leaf mottling, and narrowing. (Source: R. J. Reynolds Tobacco Company Slide Set, R.J. Reynolds Tobacco Company, Bugwood.org)

1.2.4. Tobacco Mosaic Virus (TMV)

TMV is in the top 10 of the world's most destructive plant viruses (Rybicki, 2015). TMV is a member of the *Tobamovirus* genus, which is one the largest group of plant viruses under the *Virgaviridae.* TMV has a wide host range infecting mostly Solanaceae crops, ornamental plants, and weeds (Adams *et al.,* 2017). TMV was the first-ever plant virus discovered and isolated in 1881, which opened and laid a foundation for plant virology. TMV was the first plant virus to be seen under an electron microscope, isolated from the plant material and purified and detected using electrophoresis apparatus (Kausche *et al.,* 1936; Stanely, 1935, Eriksson-Quensel and Svedberg, 1936)

TMV has rod-shaped particles of about 300nm x 15 nm in size and 2130 copies of coat protein that envelope its single-strand RNA genome (Klug, 1999). TMV RNA genome encodes for four genes, two replicase proteins directly translated from the RNA strand, and the movement protein and coat protein translated from the subgenomic RNA (Fig 1.8; Adams *et al.,* 2017). TMV has been reported in SA, and it has been a problem in many crops. The virus is easily transmissible from one host to another through direct contact, use of contaminated tools. TMV is a very stable virus, and it was reported to be infectious for 50 years stored at 4℃. TMV infection causes stunted growth with leaves showing a mosaic pattern of light and dark green and necrotic lesions (Fig 1.9a; b)

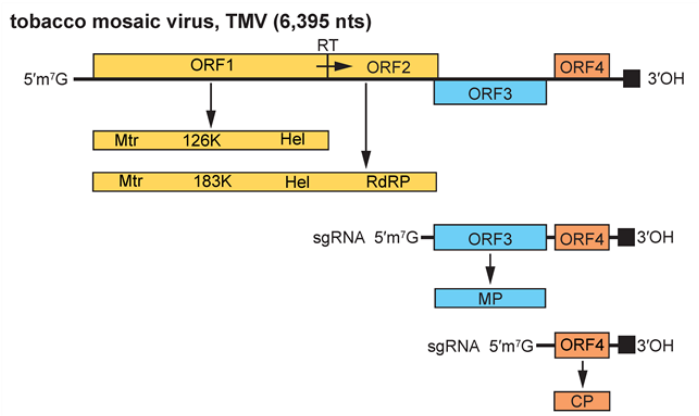


Figure 1.8: Genome organization of tobacco mosaic virus (TMV). Genomic RNA is capped, and it is a template for the expression of the 126 and 183 kDa proteins. The 3′ distal movement protein (MP) and capsid protein (CP) ORFs are expressed from separate 3′ co-terminal sgRNAs. The tRNA structure motif at the 3′-end of the RNA is represented by a dark square (ICTV, 2021).

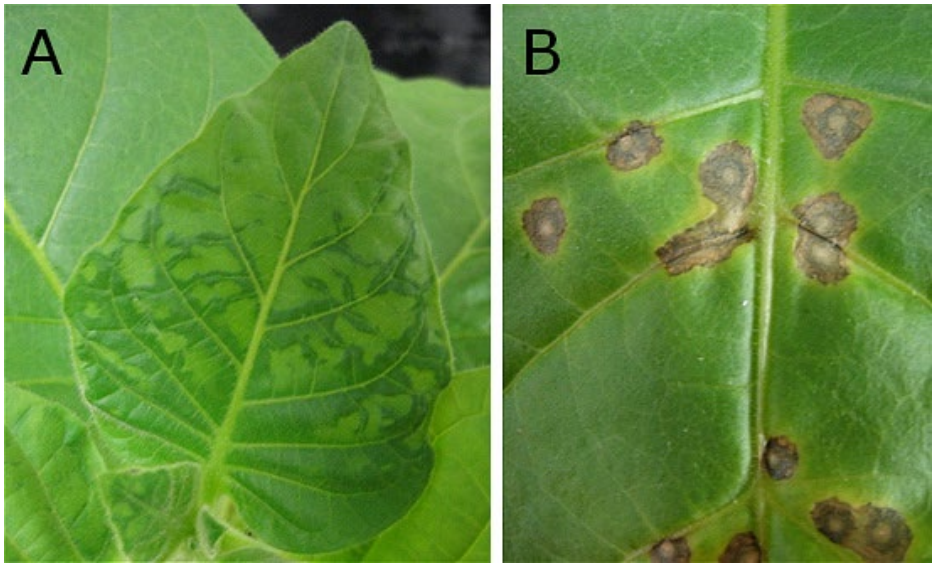


Figure 1.9.: (**a**) Systemic infections of *Nicotiana tabacum* cv. Turk plants showing TMV-associated mosaic. (**b**) Necrotic local lesions on *N. tabacum* Glurk leaf, demonstrating Holmes' N-gene resistance following inoculation with TMV. Photo: K.-B. G. Scholthof.

1.2.5. Tobacco leaf curl disease-causing viruses (TLCD)

TLCD was first reported in Netherlands East Indies in 1912; however, a disease-causing leaf curl on tobacco was already present in South Africa as early as 1902 (Paximadis *et al.,* 1999). TLCD was once reported as one of the most destructive diseases in East and Southern Africa; in Zimbabwe and Northwest provinces in SA (Paxidimidis 1997). The first few years of

research aimed to isolate geminiviruses because the literature stated that leaf curl disease was caused by whitefly-transmitted geminiviruses (Lucas, 1975; Paxidimidis 1997). There were numerous attempts to isolate the virus particles causing TLCD, but they were not successful. The virus particles resembling phytoreoviruses have consistently isolated from leaf curl-infected tobacco from the eastern and western tobacco-growing regions of South Africa (Rey et al., 1999). Later, two phytoreovuses were identified; wound tumor virus (WTV) and tobacco leaf enation virus (TLEV), in the 90s (Rey et al., 1999; Rey *et al.,* 2007). They were reported to be transmitted by grasshoppers. There were several reports on tobacco leaf curl symptoms in SA, which were suspected to be caused by geminiviruses. Tobacco leaf curl virus (TbLCV) was then discovered, and it was transmitted by whiteflies (Paximadis *et al.,* 1999). Tobacco infected with TLCD is usually curled, twisted, with enations on the inner side of the lamina along the veins, and often has stunted growth (Fig.1.10).

Figure 1.10: Tobacco plant showing typical leaf curl disease symptoms. (Jing *et al.,* 2016)

1.3. DETECTION AND CHARACTERIZATION METHODS

There are several plant virus diagnostic and detection techniques available, and more continue to be developed. Since TMV was first identified over a century ago, over 1000 plant viruses have been discovered using different detection techniques (Ashraf *et al.,* 2019). The methods are classified into symptomatology, host indexing (bioassays), serology, electron microscopy, and molecular techniques. Virus diagnosis was the preserve of specialists with years of experience in recognizing and describing virus symptoms on hosts, supplemented with complex and cumbersome methods such as bio-assays on indicator plants and the use of elaborate techniques like transmission electron microscopy (TEM) (Boonhman *et al.,* 2014). Before Enzyme-linked immunosorbent assay (ELISA) was developed. Other serological

methods were used for diagnostic purposes were immuno-specific electron microscopy (Derrick, 1976) to determine virus taxonomic relationships, e.g., Ouchterlony gel diffusion assay (Bercks, 1967).

1.3.1. Serological screening

Serological detection techniques use specific antibodies produced from rabbits, which respond to the antigen, i.e., plant virus. ELISA, Tissue blot immunoassay (TBIA), and lateral devices are currently used to detect viruses (Lima *et al.,* 2012). ELISA is based on the concept of antigen-antibody reactions; there are commercially available antibodies, and kits that can detect most viruses, including CMV, PVY, TSWV, and TMV DAS ELISA remains a highly sensitive method than lateral devices such as immunostrips and pocket diagnostics (Van Regenmortel and Dubs, 2019). Although molecular assays excel, it remains an important test because it is easy to use, time and cost-efficient, and can be used for many samples (Jeong *et al.,* 2014). The highly appropriate method for many situations, including surveillance, eradication, certification of mother plants, sanitation, and quarantine, especially when specific viruses need to be reliably detected (Van Regenmortel and Dubs, 2019).

1.3.2. Biological assays

Mechanical sap inoculation is another helpful tool for detection. The tissue from a suspected infected plant is used to inoculate a known susceptible indicator plant (Shaw, 2018). These indicator plants include *Nicotiana tabacum* L., *Nicotiana glutinosa* L., *Nicotiana benthamiana*, *Petunia X hybrida* Vilm, and *Emilia sonchifolia* (L.) (Momol *et al.,* 2004). Successful viral transmission to one of the indicator plants is helpful in two ways: it helps diagnose a plant virus through symptomatology and serves as a culture of a virus (Hull, 2009). The main problem in using bioassays is that most viruses cause similar symptoms, making it challenging to identify a particular virus (Shaw, 2018).

1.3.3. Transmission Electron microscopy (TEM)

The TEM was first developed in the 1930s by physicists in various countries, including Germany, in particular (Haguenau *et al.,* 2003). The first use of EM in clinical virology concerned the differential diagnosis of smallpox (caused by the variola virus of the poxvirus family). TEM is used to identify the viral particles' morphology and size at a genus level, and it is a good initial step for virus identification (Richert-Pöggeler., 2019). There are different ways to identify plant viruses using TEM; ultra-thin leaf sectioning, leaf dip, serological TEM,

and purified virus viewing (Griffiths and Lucocq, 2014.). TEM serves as a powerful tool to determine which detection method (i.e., bioassays, serological, or molecular approach) should subsequently be used for further virus identification (Gencer *et al.,* 2018).

1.3.4. Molecular screening

1.3.4.1. Reverse transcription-polymerase chain reaction (RT-PCR)

PCR methods for virus detection were first published in the early 1990s. It has remained the most powerful technique that allows virus detection by amplifying a particular plant viral nucleic acid present in the plant specimen (Villamor *et al.,* 2019). For RNA viruses, reverse-transcription (RT) is done to make complementary DNA (cDNA) from the viral genome in the presence of reverse transcriptase. The generated cDNA is subsequently used as a template during PCR mediated by a thermostable DNA polymerase (e.g., *Taq, Pfu, Tht*) in the presence of specific oligonucleotide primers complementary to the targeted viral nucleic acid (Jeong *et al.,* 2014). *Tth* polymerase, with its intrinsic reverse transcriptase activity, became ideal for one-step RT-PCR (Kwak *et al.,* 2020). The amplified PCR product is then analyzed using agarose gel electrophoresis. Although this technique is time-consuming and expensive, it verifies results from the serological tests (German *et al.,* 1992).

1.3.4.2. Sequencing

Next-generation sequencing (NGS) has proven to be a valuable tool for virus detection, discovery, or diversity studies and has increased in popularity while decreasing costs (Visser *et al.,* 2016). NGS, also called high-throughput sequencing (HTS), is a term used to describe modern sequencing technologies that allow the sequencing of nucleic acids much quicker, accurately, and cheaper than Sanger sequencing (Pecman *et al.,* 2017). HTS was initially introduced as next-generation sequencing to mark a distinct demarcation from the former sequencing technology commonly known as Sanger sequencing that was introduced in 1977 (Ibaba and Gubba, 2020)

As in other fields, NGS technologies have led to a revolution in virus discovery and exciting new possibilities for diagnostics. The application of massively parallel sequencing approaches, and subsequent bioinformatics analysis for viral sequences, carries the promise of routine, generic detection of viruses and other pathogens alike. Indeed several techniques were first published in 2009, applying NGS to identify diverse plant viruses, using different sequencing platforms and nucleic acid preparations as starting material (Pecman *et al.,* 2017), and many others have since followed. Whereas different sequencing platforms have been used

(principally Roche 454 and Illumina), which can make a difference in cost, ease of sequence assembly, and identification, the key distinctions lie in the nucleic acid purification techniques employed (Krishna *et al.,* 2019).

Table 1.1: Comparison of the main features of the popular next-generation sequencing platforms (Krishna *et al.,* 2019)

Platform	Maximum throughput Mb/run	Mean length (nt)	Error rate	Specific features	Origin of errors
Roche/454	700	Up to 1000 bp	10^{-3}–10^{-4}	Long read lengths (improved mapping in repetitive regions, short runtime	Homopolymers, intensity cut-off, signal cross-talk interference among neighbors, amplification, mixed beads
Ion Torrent PGM	1000	~200	3×10^{-2}	Stable sequence quality, better sequencing GC depth distribution	Homopolymers, amplification
Illumina	6000	~100	10^{-2} -10^{-3}	Highest throughput Long-/short-run times, low capital cost, low-cost per Mb	Signal interference among neighboring clusters, homopolymers, phasing, nucleotide labeling, amplification, low coverage of AT-rich regions
SOLiD	20,000	~50	10^{-2}–10^{-3}	High throughput, highest accuracy two-base encoding provides inherent error correction	Intensity cut-off, homopolymers, signal cross-talk interference among neighbors, amplification, mixed beads

1.3.4.2.1. High-throughput RNA Sequencing procedure

Figure 1.7 shows the summary of how High-throughput RNA sequencing is done. The frozen tissue samples are ground in the lysis buffer, and the total RNA is isolated from the host cell using an RNA extraction kit or other protocols. The mRNA produces the first and second cDNA strand, and the double-stranded cDNA molecules are then fragmented using specific enzymes (Krishna *et al.,* 2019). TA ligation of the barcoded adapters is facilitated by repairing the ends of these molecules and adding an A nucleotide. The ligated samples are then enriched by amplification using adapter-specific primers and purified for sequencing (Sudhagar *et al.,* 2018).

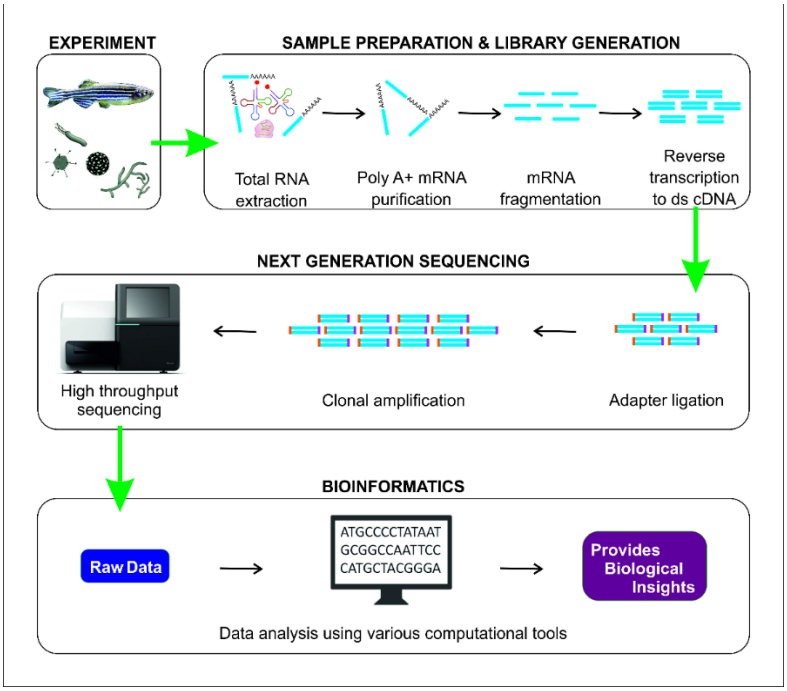


Figure 1.7: Outline of the high-throughput RNA-seq (HTR) steps from library preparation to NGS then bioinformatics (Sudhagar *et al.,* 2018)

1.4. MANAGEMENT STRATEGIES FOR PLANT VIRUSES

Managing plant viral diseases is challenging because of the complex and dynamic nature of virus epidemics and their evolution (Acosta-Leal *et al.,* 2011; Elena *et al.,* 2014). Once the disease has entered the plant, the host-pathogen relationship is compatible, and it is practically impossible to break it or reduce disease development (Momol *et al.,* 2004). It is crucial to consider the genetic diversity and virus populations for efficient and sustainable control strategies and efficient, fast, and reliable identification tools. The control strategies need to be practiced using integrated pest management (IPM), including phytosanitation, cultural, host-

plant resistance, chemical, and biological control measures (Gent *et al.,* 2006). Disease management of plant viruses is also is based on two approaches: host plant resistance to viral infections and prophylactic measures to reduce the spread of viruses (Rubio *et al.,* 2020).

1.4.1. Phytosanitation

Phytosanitation minimizes or eliminates virus contamination from infected plant propagules, such as infected seeds, roots, tubers, cuttings, and infected rootstocks grafted with healthy scions or reservoirs of infection infected volunteer crop plants or weeds (Naik and Buckseth, 2018). Phytosanitary measures include the removal of alternative viruses and vector hosts. Virus elimination includes planting healthy certified seeds, seedlings, and vegetative propagules, roguing out symptomatic plants (Rubio *et al.,* 2020). Virus and vector elimination include disinfecting the planting area and tools used, soil solarization, compost sterilization, fallowing, UV nutrient solution irradiation, filtered irrigation water. Phytosanitary measures can reduce internal or external virus sources or both together, but they are effective only early stages of plant development (Maree *et al.,* 2018)

1.4.2. Cultural and physical control

Cultural control measures include destroying the old infected crop remains, alternative hosts, and removal of weed hosts so that they will not act as a "green bridge" for infections. Crop rotation and intercropping can be methods that can reduce disease epidemics (Gent *et al.,* 2006). Cultural measures include manipulation of planting or harvesting date to avoid peak times of insect vectors; increased sowing depth or row spacing to avoid or minimize virus spread by insect vectors, planting non-host crops as barriers around the field and in between virus infection sources and a susceptible crop, use of protective row covers and nets or reflective or straw mulches against insect vectors, and altering soil pH or irrigation water to minimize nematode, fungi or protist vector activity in contaminated soil (Naik and Buckseth, 2018).

Fine mesh screens on doors and windows are a reliable solution for greenhouse crops among the physical vector control methods. Some insects are color sensitive; therefore, they respond negatively to UV-reflective mulches (Momol *et al.,* 2004). The use of aluminum-surfaced plastic film mulch reduces the immigration of thrips, whiteflies, aphids, and the spread of viral diseases. The use of yellow or blue sticky plates for the early detection of insect vectors in greenhouses is also helpful, as it allows for the rapid adoption of control measures (Rosello *et*

al., 1996). Cultural and physical control strategies can reduce the virus sources in all stages, early, mid-and later stages of plant growth (Jones and Naidu, 2019).

1.4.3. Chemical control

Although insecticides and biocontrol strategies exist for insect vectors, there are no reliable chemical or biological strategies to control plant viruses (Jones and Naidu, 2019). Chemical control of vectors is highly effective and can be promoted. Although chemicals may be helpful to reduce the virus spread, the virus mode of transmission must be well understood so that proper chemicals and application methods can be used (Higley, 2019). There is also potential for insect repellents for vector management; for example, some work has been done with visual repellents (mulches, paints, colored netting). Chemical control includes proper insecticides, the number and frequency of applications, the application method (spray particle size), and pesticide rotation (Riley and Pappu, 2004).

1.4.4. Biological control

The use of natural vector predators, such as bacterial toxins or botanical compounds, can help control vector populations (Altieri *et al.,* 2018). Biological control of the pathogen is another important option. There are a few examples of biological control strategies; the use of fungi against the vector, predatory bugs, and mites (Maree *et al.,* 2018). The Deuteromycetes fungal species that can be used as a bioinsecticide to reduce thrips population are *Beauveria bassiana* and *Metarhizium anisopliae*, and there are also the predatory bugs from genus *Orius* and *Amblyseius cucumeris* mite that feeds on thrips (Woldemelak, 2020). Plant growth-promoting fungi have been reported to induce systemic resistance against CMV, and these fungi are *Trichoderma asperellum* SKT-1, *Glomus mosseae*, *Fusarium equiseti,* and *Penicillium simplicissimum* GP17-2. (Collatz *et al.,* 2021).

1.4.5. Host plant resistance

The resistant host cultivars may occur through natural breeding, and they can also be genetically engineered. Genetically resistant cultivars are effective viral disease control against uncontrollable vector populations (Lapidot *et al.,* 2014). The use of resistant crop cultivars helps to protect the environment from harmful chemicals. Breeding for resistance is one of the challenging methods because it takes time to find resistant genotypes, and it is expensive (Pechinger *et al.,* 2019). Several factors are considered by plant pathologists and plant breeders

when breeding for resistance; they have to screen for resistance while also assessing the phenotype (García-Arenal and McDonald, 2003).

The aim of breeding to get the complete resistance in which a virus cannot start any systemic infection, which is done in different approaches. Gene editing is one of the methods to mutagenize the host genes involved in recessive resistance to prevent interaction with viruses (Chandrasekaran *et al.,* 2016). Another strategy is to develop resistant cultivars using the RNA silencing mechanism, a natural antiviral defense mechanism (Kaldis *et al.,* 2018). Transgenic plants with multiple constructs from different viral genomes (from the same species and different species) can be used to minimize the risk of resistance breakdown (Duan *et al.,* 2012). Another method for viral disease prevention used in different viruses is cross-protection. It is based on inoculating mild or attenuated viral strains to protect plants against severe strains of the same virus (Pechinger *et al.,* 2019). In order to use cross-protection, it is important to evaluate the genetic and biological variability of the virus population and search for the mild strain that is closely related to the severe ones (Hanssen *et al.,* 2010).

1.5. CONCLUSION

In conclusion, identifying viruses has been challenging, and some attempts to isolate and identify viruses were unsuccessful. These difficulties are that viruses may elicit similar symptoms on the host and have the same biological properties. One vector can transmit more than one virus, making it challenging to identify the viruses using specific methods like RT-PCR and Serological methods. The symptoms of some plant viruses may be confused with those caused by high temperature, insect feeding, growth regulators, herbicides, mineral deficiencies, and mineral excesses. Plant viral diseases cannot be diagnosed based on symptoms alone. Therefore, much work needs to be done to achieve more efficient and reliable identification methods.

Fortunately, over the past 20 years, NGS, currently known as HTS, has been introduced as the more effective and reliable method to identify viruses without prior knowledge or any visible symptoms and vector presence. The HTS technology is also to identify more than one virus simultaneously and discover new emerging viruses. The use of more accurate and efficient methods for virus detection makes it easier to understand more about virus genetic diversity, which is important for developing sustainable strategies to manage viral diseases. Previous studies in SA were based on specific methods for plant virus isolation and identification, which made it challenging to have an accurate and time-efficient virus detection.

1.6. REFERENCES

Acosta-Leal, R., Duffy, S., Xiong, Z., Hammond, R.W. and Elena, S.F., 2011. Advances in plant virus evolution: translating evolutionary insights into better disease management. *Phytopathology*, *101*(10):1136-1148.

Adams, M.J., Adkins, S., Bragard, C., Gilmer, D., Li, D., MacFarlane, S.A., Wong, S.M., Melcher, U., Ratti, C., Ryu, K.H. and Consortium, IR, 2017. ICTV virus taxonomy profile: Virgaviridae. *The Journal of General Virology*, *98*(8): 1999.

AGRISETAhttps://www.agriseta.co.za/wpcontent/uploads/2021/02/Agriseta_Tobacco_SSSP_DIGITAL.pdf

Altieri, M.A., Farrell, J.G., Hecht, S.B., Liebman, M., Magdoff, F., Murphy, B., Norgaard, R.B. and Sikor, T.O., 2018. Plant disease ecology and management. In *Agroecology* : 307-319

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.

Ashraf, S., Saeed, S., Musheer, N., Choudhary, A. and Jamil, A., 2019. Role of molecular techniques, transgenic plants and microorganisms for managing viral diseases of plants. *Management*, *201*:223.

Bellstedt, D.U., Glais, L., Davie, K. and Lacomme, C., 2017. Evolution and Origin of PVY. In Potato virus Y*: Biodiversity, Pathogenicity, Epidemiology And Management*:77-101

Bercks, R., 1967. Methodische Untersuchungen über den serologischen Nachweis pflanzenpathogener Viren mit dem Bentonit-Flockungstest, dem Latex-Test und dem Bariumsulfat-Test. *Journal of Phytopathology*, *58*(1):1-17.

Best, RJ, 1968. Tomato spotted wilt virus. In *Advances in virus research* (Vol. 13, pp. 65-146). Academic Press.Best RJ (1968). Tomato spotted wilt virus. *Advances in Virus Research* 13:65145

Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J. and Mumford, R., 2014. Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Research*, *186*:20-31.

Brittlebank, C.C., 1919. Tomato diseases. *Journal of the Department of Agriculture in Victoria*. 17:231–235.

Canto, T., Prior, D.A., Hellwald, K.H., Oparka, K.J. and Palukaitis, P., 1997. Characterization of Cucumber mosaic virus. *Virology*, *237*(2),:237-248.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T. and Gal-On, A., 2016. Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Molecular Plant Pathology*, *17*(7):1140-1153.

Christenhusz, M.J., Fay, M.F. and Chase, M.W., 2017. *Plants of the world*. University of Chicago Press, Chicago, Illinois, United States.

Collatz, J., Hinz, H., Kaser, J.M. and Freimoser, F.M., 2021. Benefits and risks of biological control. *Biological Control: Global Impacts, Challenges and Future Directions of Pest Management*: 142-165.

Derrick, K.S. and Brlansky, R.H., 1976. Assay for viruses and mycoplasmas using serologically specific electron microscopy. *Phytopathology*, *66*(6):815-820.

Dombrovsky, A. and Smith, E., 2017. Seed transmission of Tobamoviruses: Aspects of global disease distribution. *Advances in Seed Biology*:233-260.

Duan, C.G., Wang, C.H. and Guo, H.S., 2012. Application of RNA silencing to plant disease resistance. *Silence*, *3*(1):1-8.

Elena, S.F., Fraile, A. and García-Arenal, F., 2014. Evolution and emergence of plant viruses. *Advances in Virus Research*, *88*:161-191.

Eriksson-Quensel, I.B. and Svedberg, T., 1936. Sedimentation and electrophoresis of the tobacco-mosaic virus protein. *Journal of the American Chemical Society*, *58*(10), pp.1863-1867.

FAOSTAT <http://www.fao.org/faostat/en/#data/QV accessed 26 September 2021

Flood, J., 2010. The importance of plant health to food security. *Food Security*, *2*(3): 215-231.

Garcıa-Arenal, F. and Palukaitis, P., 2009. Cucumber mosaic virus. *Desk Encyclopedia of Plant and Fungal Virology, Elsevier professional: sl*:171-6.

Gencer, D., Bayramoglu, Z., Nalcacioglu, R., Kleespies, R.G., Demirbag, Z. and Demir, I., 2018. Characterisation of three Alphabaculovirus isolates from the gypsy moth, *Lymantria dispar* \ (Lepidoptera: Erebidae), in Turkey. *Biocontrol Science and Technology*, *28*(2):107-121.

Gent, D.H., du Toit, L.J., Fichtner, S.F., Krishna Mohan, S., Pappu, H.R. and Schwartz, H.F., 2006. Iris yellow spot virus: an emerging threat to onion bulb and seed production. *Plant Diseases.* 90:1468–1480.

German, T.L., Ullman, D.E. and Moyer, J.W., 1992. *Tospoviruses*: diagnosis, molecular biology, phylogeny, and vector relationships. *Annual Review in Phytopathology* 30: 315–348.

Glais, L., Bellstedt, D.U. and Lacomme, C., 2017. Diversity, characterisation and classification of PVY. In *Potato virus Y: Biodiversity, Pathogenicity, Epidemiology And Management*:43-76

Griffiths, G. and Lucocq, J.M., 2014. Antibodies for immunolabeling by light and electron microscopy: not for the faint hearted. *Histochemistry and Cell Biology*, *142*(4):347-360.

Grisan, S., Polizzotto, R., Raiola, P., Cristiani, S., Ventura, F., Di Lucia, F., Zuin, M., Tommasini, S., Morbidelli, R., Damiani, F. and Pupilli, F., 2016. Alternative use of tobacco as a sustainable crop for seed oil, biofuel, and biomass. *Agronomy for Sustainable Development*, *36*(4): 1-8.

Guo, Y., Jia, M.A., Yang, Y., Zhan, L., Cheng, X., Cai, J., Zhang, J., Yang, J., Liu, T., Fu, Q. and Zhao, J., 2017. Integrated analysis of tobacco miRNA and mRNA expression profiles under PVY infection provids insight into tobacco-PVY interactions. *Scientific Reports*, *7*(1):1-10.

Haguenau, F., Hawkes, P.W., Hutchison, J.L., Satiat–Jeunemaître, B., Simon, G.T. and Williams, D.B., 2003. Key events in the history of electron microscopy. *Microscopy And Microanalysis,* 9(2):96-138

Hančinský, R., Mihálik, D., Mrkvová, M., Candresse, T. and Glasa, M., 2020. Plant viruses infecting Solanaceae family members in the cultivated and wild environments: A Review. *Plants*, *9*(5):667.

Hanssen, I.M., Gutiérrez-Aguirre, I., Paeleman, A., Goen, K., Wittemans, L., Lievens, B., Vanachter, A.C., Ravnikar, M. and Thomma, B.P.H.J., 2010. Cross-protection or enhanced

symptom display in greenhouse tomato co-infected with different Pepino mosaic virus isolates. *Plant Pathology*, *59*(1): 13-21.

Higley, L., 2019. *Integrated management of insect pests: Current and future developments*. Burleigh Dodds Science Publishing. Cambridge, United Kingdom

Hull, R., 2009. *Comparative plant virology*. Academic press. Norwich United Kingdom

Ibaba, J.D. and Gubba, A., 2020. High-throughput sequencing application in the diagnosis and discovery of plant-infecting viruses in Africa, a decade later. *Plants*, *9*(10): 1376.

ICTV online: https://talk.ictvonline.org/taxonomy/ accessed 05 May 2021

Jeong, J.J., Ju, H.J. and Noh, J., 2014. A review of detection methods for the plant viruses. *Research in Plant Disease*, *20*(3): 173-181.

Jing, C., Wang, C., Li, K., Wu, G., Sun, X. and Qing, L., 2016. Molecular identification of tobacco leaf curl disease in Sichuan province of China. *Virology Journal*, *13*(1): 1-5.

Jones, R.A. and Naidu, R.A., 2019. Global dimensions of plant virus diseases: Current status and future perspectives. *Annual Review of Virology*, *6*: 387-409

Kaldis, A., Berbati, M., Melita, O., Reppa, C., Holeva, M., Otten, P. and Voloudakis, A., 2018. Exogenously applied dsRNA molecules deriving from the Zucchini yellow mosaic virus (ZYMV) genome move systemically and protect cucurbits against ZYMV. *Molecular Plant Pathology*, *19*(4): 883-895.

Kausche, G.A., Pfankuch, E. and Ruska, H., 1939. Die Sichtbarmachung von pflanzlichem virus im Übermikroskop. *Naturwissenschaften*, *27*(18): 292-299.

Klug, A., 1999. The tobacco mosaic virus particle: structure and assembly. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *354*(1383):531-535.

Korbecka-Glinka, G., Przybyś, M. and Feledyn-Szewczyk, B., 2021. A Survey of Five Plant Viruses in Weeds and Tobacco in Poland. *Agronomy*, *11*(8):1667.

Kormelink, R., van Poelwijk, F., Peters D., and Goldbach, R., 1992. Non-viral heterogeneous sequences at the 5' ends of tomato spotted wilt virus mRNAs. *Journal of General Virology 73: 2125-2128*

Krishna, B.M., Khan, M.A. and Khan, S.T., 2019. Next-generation sequencing (NGS) platforms: An exciting era of genome sequence analysis. In *Microbial Genomics in Sustainable Agroecosystems:*89-109

Kwak, H.R., Go, W.R., Hong, S.B., Kim, J.E., Kim, M. and Choi, H.S., 2020. One-step multiplex RT-PCR for simultaneous detection of four viruses infecting Rehmannia glutinosa. *Journal of General Plant Pathology*, *86*(2): 143-148.

Lapidot, M., Legg, J.P., Wintermantel, W.M. and Polston, J.E., 2014. Management of whitefly-transmitted viruses in open-field production systems. *Advances in Virus Research*, *90*:147-206.

Lima, J.A.A., Nascimento, A.K.Q., Radaelli, P. and Purcifull, D.E., 2012. Serology applied to plant virology. *Serological Diagnosis of Certain Human, Animal and Plant Diseases. Rijeka, Croatia: Intech*, *1*: 71-94.

Liu, D., Zhao, Q., Cheng, Y., Li, D., Jiang, C., Cheng, L., Wang, Y. and Yang, A., 2019. Transcriptome analysis of two cultivars of tobacco in response to Cucumber mosaic virus infection. *Scientific Reports*, *9*(1): 1-12.

Lucas, G.B., 1975. "Diseases of tobacco". *Diseases of Tobacco*

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.

Mishra, R., Verma, R.K., Sharma, P., Choudhary, D.K. and Gaur, R.K., 2014. Interaction between viral proteins with the transmission of Potyvirus. *Archives of Phytopathology and Plant Protection*, *47*(2): 240-253.

Momol, M.T., Olson, S.M., Funderburk, J.E., Stavisky, J. and Marois, J.J., 2004. Integrated management of tomato spotted wilt on field-grown tomatoes. *Plant Diseases* 88: 882–890.

Moore, E.S. and Andessen, E.E., 1939. Notes on Plant Virus Diseases in South Africa. I. The Kromnek Disease of Tobacco and Tomato. Dept. Agric., Union of South Africa. *Scientific Bull*, *182*: 1-36.

Moore, E.S., 1932. A virus disease of tobacco in South Africa. *Nature*, *129*(3258): 544-544.

Naik, P.S. and Buckseth, T., 2018. Recent advances in virus elimination and tissue culture for quality potato seed production. *Biotechnologies of Crop Improvement, Volume 1*: 131-158.

Nemes, K., Gellért, Á., Bóka, K., Vági, P. and Salánki, K., 2019. Symptom recovery is affected by Cucumber mosaic virus coat protein phosphorylation. *Virology*, *536*: 68-77.

Pammel, L.H., 1911. *A manual of poisonous plants, chiefly of eastern North America: With brief notes on economic and medicinal plants, and numerous illustrations*. Torch Press.Cedar Rapida, Iowa

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.

Pavithra, B.S., Krishnareddy, M. and Rangaswamy, K.T., 2016. Detection and partial characterization of groundnut bud necrosis virus in chilli. *International Journal of Natural Science 7*(4):843-847.

Paximadis, M., Dusterwald, G., Duyver, L. and Rey, M.E.C., 1997. Preliminary investigation of the causal agent (s) of a disease-causing leaf curl of tobacco in South Africa. *Plant Pathology*, *46*(2): 282-290.

Paximadis, M., Idris, A.M., Torres-Jerez, I., Villarreal, A., Rey, M.E.C. and Brown, J.K., 1999. Characterization of tobacco geminiviruses in the Old and New World. *Archives of Virology*, *144*(4): 703-717.

Pechinger, K., Chooi, K.M., MacDiarmid, R.M., Harper, S.J. and Ziebell, H., 2019. A new era for mild strain cross-protection. *Viruses*, *11*(7), p.670.

Pecman, A., Kutnjak, D., Gutiérrez-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.

Qiu, Y., Zhang, Y., Hu, F. and Zhu, S., 2017. Characterization of siRNAs derived from Cucumber mosaic virus in infected tobacco plants. *Archives of Virology*, *162*(7): 2077-2082.

Randall RP, 2012. A Global Compendium of Weeds. Perth, Australia: Department of Agriculture and Food Western Australia: 1124 http://www.cabi.org/isc/FullTextPDF/2013/20133109119.pdf

Rey, C., 2020. The hunt for plant viruses in South Africa: The early days. *South African Journal of Science*, *116*(11-12), 1-8.

Rey, M.E.C., D'Andrea, E., Calvert-Evers, J., Paximadis, M. and Boccardo, G., 1999. Evidence for a phytoreovirus associated with tobacco exhibiting leaf curl symptoms in South Africa. *Phytopathology*, *89*(4): 303-307.

Richert-Pöggeler, K.R., Franzke, K., Hipp, K. and Kleespies, R.G., 2019. Electron microscopy methods for virus diagnosis and high-resolution analysis of viruses. *Frontiers in Microbiology*, *9*: 3255.

Riley, D.G. and Pappu, HR, 2004. Tactics for management of thrips (*Thysanoptera: Thripidae*) and tomato spotted wilt virus in tomato. *Journal Of Economic Entomology*, 97,: 1648–1658.

Roselló, S., Díez, M.J. and Nuez, F., 1996. Viral diseases causing the greatest economic losses to tomato crops. I. The tomato spotted wilt virus: a review. *Scientia Horticulturae*,*67*, :117–150

Ross, H., 1986. Potato breeding–problems and perspectives: advances in plant breeding. *Journal of Plant Breeding*, *13*.

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*: 1092.

Rudolph, C., Schreier, P.H. and Uhrig, J.F., 2003. Peptide-mediated broad-spectrum plant resistance to tospoviruses. *Proceedings of the National Academy of Sciences*, *100*(8): 4429-4434.

Rybicki, E.P., 2015. A Top Ten list for economically important plant viruses. *Archives of Virology*, *160*(1): 17-20.

Samuel, G., 1930. Investigations on 'spotted wilt' of Tomatoes in Australia. *Commonwealth of Australia. Council Science and Industrial Research Bull*, *44*: 8-11.

Scholthof, K.B.G., 2004. Tobacco mosaic virus: a model system for plant biology. *Annual; Review of Phytopathology*, *42*: 13-34.

Shaw, J.G., 2018. Early events in plant virus infections. In *Molecular Plant Virology*: 1-21

Sivparsad, B.J. and Gubba, A., 2008. Isolation and molecular characterization of Tomato spotted wilt virus (TSWV) isolates occurring in South Africa. *African Journal of Agricultural Research*, *3*(6): 428-434

Stanley, W.M., 1935. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus. *Science*, *81*(2113): 644-645.

Status of Tobacco Production and Trade in Africa – Factsheets (August 2020) NTM code Source: Consolidated Tariff Schedules database (CTS), http:\\tdf.wto.org Source: Most recent national survey available to WHO: 2019 ; STEPS Survey 375 16 % of agricultural land Source: http://wits.worldbank.org/ accessed 06 August 2021

Sudhagar, A., Kumar, G. and El-Matbouli, M., 2018. Transcriptome analysis based on RNA-Seq in understanding pathogenic mechanisms of diseases and the immune system of fish: a comprehensive review. *International Journal of Molecular Sciences*, *19*(1): 245.

Swanepoel, A. and Nel, L., 1995. Epidemiological aspects of cucumber mosaic virus on flue-cured tobacco in South Africa. *African Plant Protection*, *1*(1): 13-18.

Thompson, G.J. and Van Zijl, J.J.B., 1995. Control of tomato spotted wilt virus in tomatoes in South Africa. *Tospoviruses and Thrips of Floral and Vegetable Crops 431*: 379-384.

Tungadi, T., Donnelly, R., Qing, L., Iqbal, J., Murphy, A.M., Pate, A.E., Cunniffe, N.J. and Carr, J.P., 2020. Cucumber mosaic virus 2b proteins inhibit virus-induced aphid resistance in tobacco. *Molecular Plant Pathology*, *21*(2): 250-257.

Turina, M., Kormelink, R. and Resende, R.O., 2016. Resistance to tospoviruses in vegetable crops: epidemiological and molecular aspects. *Annual Review of Phytopathology*, *54*: 347-371.

Turina, M., Tavella, L. and Ciuffo, M., 2012. Tospoviruses in the Mediterranean area. *Advances in virus research*, *84*: 403-437.Van Regenmortel, M.H.V. and Dubs, M.C., 2019. Serological procedures. *Diagnosis of Plant Virus Diseases*: 159-214.

Vijayapalani, P., Maeshima, M., Nagasaki-Takekuchi, N. and Miller, W.A., 2012. Interaction of the trans-frame potyvirus protein P3N-PIPO with host protein PCaP1 facilitates potyvirus movement. *PLoS Pathogens*, *8*(4): 1002639.

Visser, M., Bester, R., Burger, J.T. and Maree, H.J., 2016. Next-generation sequencing for virus detection: covering all the bases. *Virology Journal*, *13*(1): 1-6.

Whitfield, A.E., Ullman, D.E. and German, T.L., 2005. Tospovirus-thrips interactions. *Annual Review of Phytopathology* 43: 459–489

Wijkamp, I., WETERING, F.V.D., Goldbach, R. and Peters, D., 1996. Transmission of tomato spotted wilt virus by Frankliniella occidentalism median acquisition and inoculation access period. *Annals of Applied Biology*, *129*(2): 303-313.

Woldemelak, W.A., 2020. The major biological approaches in the integrated pest management of onion thrips, *Thrips tabaci* (Thysanoptera: Thripidae). *Journal of Horticultural Research*, *28*(1): 13-20.

Wylie, S.J., Adams, M., Chalam, C., Kreuze, J., López-Moya, J.J., Ohshima, K., Praveen, S., Rabenstein, F., Stenger, D., Wang, A. and Zerbini, F.M., 2017. ICTV virus taxonomy profile: Potyviridae. *The Journal of General Virology*, *98*(3): 352.

Chapter 2: Identification and characterization of Orthotospoviruses infecting tobacco (*Nicotiana* sp.) in South Africa

ABSTRACT

Orthotospoviruses are a group of viruses that have been reported as some of the most destructive plant viruses worldwide due to the economic damage they cause on several crops and their efficient transmission by different thrips species. The most common and widely distributed orthotospoviruses are tomato spotted wilt virus (TSWV), groundnut ringspot virus (GRSV), and tomato chlorotic spot virus (TCSV). The study aimed to identify the orthotospoviruses infecting tobacco in South Africa (SA) based on the fact that during the 2018/19 growing season, symptoms typical of orthotospovirus infection were observed in tobacco (*Nicotiana tabacum* L.) fields across SA. Large thrips populations accompanied these symptoms. Leaf samples showing typical orthotospovirus symptoms were collected from farms in three major tobacco-producing provinces. Samples were then tested for orthotospoviruses using, double antibody sandwich-Enzyme-linked immunoassay (DAS-ELISA), biological assays, Reverse transcription-polymerase chain reaction (RT-PCR), and Sanger sequencing. The obtained sequences were subsequently subjected to the Basic Local Alignment Search Tool (BLASTn). The sequences of virus isolates from the different SA provinces (obtained from this study), and sequences of previously identified isolates obtained from GenBank were selected and aligned using Clusta*W* for phylogenetic analysis based on the best fit model proposed by MEGA X software. The ELISA and RT-PCR results showed that 19 of 22 samples were positive when tested using orthotospovirus specific antibodies and orthotospovirus generic primers. Mechanical inoculation onto *Nicotiana tabacum* cv. Xanthi nc and *Nicotiana benthamiana* produced typical orthotospovirus symptoms, including vein clearing, chlorotic spots, and mosaics. The results showed that there are two orthotospoviruses infecting tobacco in SA; GRSV and TSWV. TSWV was found in all three tobacco-growing provinces, and GRSV was found in Limpopo and Northwest. To our knowledge, this is the first report of GRSV infecting tobacco in SA since it was first reported on groundnut in 1966. The information generated from this study will be useful in devising effective strategies to manage viral diseases of tobacco in SA.

2.1. INTRODUCTION

Tobacco is the most significant annual cash crop in South Africa, with over 30 million kilograms produced annually, and contributes over R17 billion to the national fiscus every year (FAO, 2021). In SA, a small farmer can earn a net income of R5000/ha with tobacco cultivation, higher than other crops such as maize (WHO, 2019). Tobacco is most commonly used as a cigarette worldwide. Recently, research has shown the potential of tobacco as a biofuel. To this end, tobacco can produce up to three times ethanol per acre as corn and three times the oil per acre as soybean (Grisan *et al.,* 2016).

Several emerging and recurrent plant viruses negatively affect tobacco production, yield, and quality worldwide (Dong *et al.,* 2010). Orthotospoviruses are economically important viruses that infect *Solanaceae* crops, especially tobacco, causing severe leaf damage, thus decreasing its yield and marketable value (Oliver & Whitfield, 2016). Symptoms on infected plants include the following: chlorotic or necrotic local lesions, ringspot, concentric ringspot, line patterns, green island mosaic on leaves, stem discoloration, followed by systemic symptoms such as wilting, stunting, mottling, crinkling, bronzing, distortion (curling), rugosity, chlorosis and (top) necrosis, which may vary on the same host species (EPPO, 2020)

According to the International Committee on Taxonomy of Viruses (ICTV), orthotospoviruses belong to the *Tospoviridae* family. Orthotospovirus virus particles are double enveloped spherical, with a diameter of 80-120 nm, and have glycoprotein projections (Turina *et al.,* 2016). Orthotospoviruses consist of a negative-sense RNA tripartite genome, consisting of Small (S), medium (M), and large (L) segments. Orthotospoviruses use an ambisense replication strategy to produce proteins (Kaye et al., 2011). The L RNA encodes the RNA-dependent RNA polymerase in a negative sense (de Haan *et al.,* 1990; Kormelink *et al.,* 1992). The M RNA encodes two proteins: the non-structural movement protein (NSm) in the positive sense and the Gn-Gc glycoprotein in the negative sense (de Haan *et al.,* 1990). The S RNA encodes the non-structural protein (NSs) in the positive sense and the nucleocapsid (N) protein in the negative sense (Whitfield *et al*., 2005; Pappu *et al*., 2009).

In SA, tomato spotted wilt virus (TSWV) was the first Orthotospovirus reported on tobacco in 1905 and was known as ‘wilt of tobacco,’ which was later named ’Kromnek’ (Moore, 1933). To date, no other orthotospovirus has been reported on tobacco in SA. This study looked at the prevalence of Orthotospoviruses across different tobacco growing regions in SA based on the symptoms and thrips populations observed. The study outlines protocols for identifying and

characterizing different orthotospoviruses and assesses their phylogenetic relationship to those viruses that occur in other geographical regions.

2.2. MATERIALS AND METHODS

2.2.1. Field survey and sampling

In the 2019 growing season, significant thrips populations were observed in tobacco fields countrywide. The thrips populations were accompanied by typical Orthotospovirus symptoms, including severe necrosis, yellowing, stunted growth symptoms, and leaf curling (Fig. 2.1 A-D). The leaf material exhibiting typical Orthotospovirus symptoms in the presence of thrip vectors was collected from the major tobacco-growing regions shown in the map (Table 2.1; Fig 2.2). Six different tobacco cultivars were used as sources of the samples used in this study (Table 2.1). Samples were placed in the 50ml labeled bottles and shipped in dry ice to the Plant Virology laboratory, Pietermaritzburg Agriculture campus, the University of KwaZulu-Natal. They were stored in a -80℃ freezer until analyzed. A total of 22 samples were used for this study.

2.2.2. Biological assays

Mechanical inoculations were done at the University of KwaZulu-Natal Pietermaritzburg Controlled Environmental Unit (CERU; 29°37'32.9" S 30°24'18.8" E). Disease-free tobacco *(Nicotiana tabacum cv. Xanthi nc. and Nicotiana benthamiana*) seedlings were grown in 15cm pots at 25℃ and 85% Relative humidity in a growth room with constant light intensity. Plants at the 4-leaf stage were inoculated with the infected field samples prepared with an ice-cold phosphate buffer (pH 7.5). Leaves were dusted with a carborundum powder, gentley rubbed with the sap solution using the pestle, and rinsed after 2 minutes with distilled water. The inoculated plants were monitored every day for symptom development over four weeks. In total, 46 pot plants were inoculated, with each sample being inoculated twice. There were two controls; one was mock inoculated with phosphate buffer and the other with distilled water.

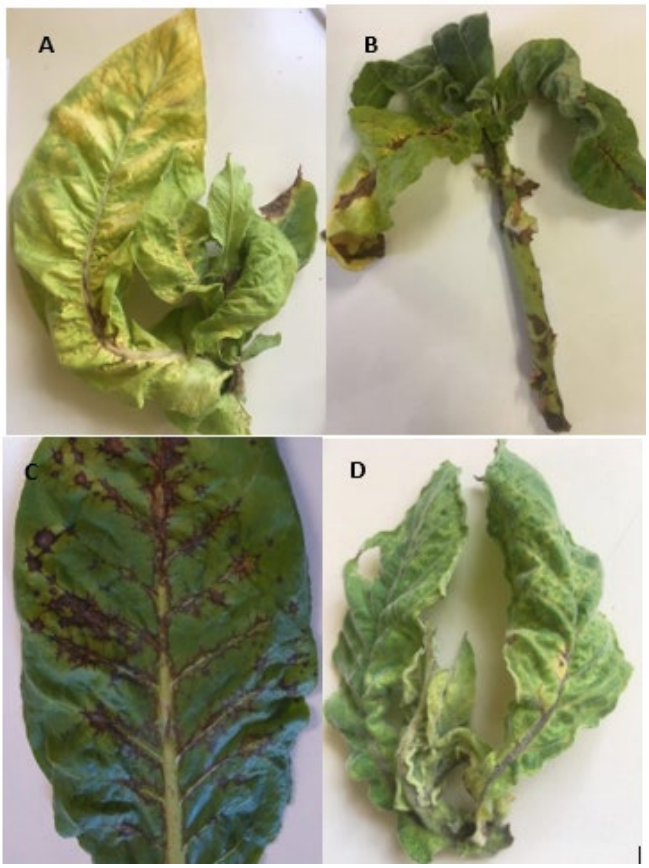


Figure 2.1: Orthotospovirus associated symptoms observed on tobacco plants during the field survey from major tobacco growing areas Rustenburg, Outdshroon, Globlerdal, and Ellisras. A and C: interveinal chlorosis with necrosis; B: stem necrosis and stunted growth; D: Leaf curling with chlorotic spots and interveinal necrosis

2.2.3. Double sandwich Enzyme-linked immunoassay (DAS-ELISA)

The presence of orthotospoviruses was detected using the double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) with commercially available orthotospovirus antibodies (Bio-Rad, USA) according to the manufacturer's instructions. Both the coating and the conjugate antibodies were diluted by a factor of 1/1000 in their respective buffers. About 0.5g of the leaf material for each sample was ground in a 2ml Eppendorf tube with liquid nitrogen using a sterile drill bit. To each ground leaf sample, 1.5ml of general extraction buffer [PBS-Tween 20 with 2% (w/v) PVP-40 pH 7.4] was added and mixed by vortexing. The homogenate was then centrifuged at 13000*xg* for 3 minutes.

A volume of 100µl was used in each well for all reagents added to the plate. Plates were first coated with coating antibody followed by incubation at 37C for two hours. After incubation, plates were washed with washing buffer to remove any unattached antibodies. Prepared

samples were then added to the plate in duplicate. Positive control with orthotospovirus antigen from the kit was included. The disease-free leaf sap from uninoculated and buffer inoculated plants was used as negative and buffer controls, respectively.

The ELISA plate was then incubated overnight at 4℃, after which plates were washed with washing buffer. This was followed by adding the antibody-enzyme conjugate, after which plates were incubated at 37℃ for 2 hours. This was followed by plate washing. The final step was the addition of 1mg/ml p-nitro phenyl phosphate (pNPP) liquid substrate. Plates were then incubated in the dark room temperature until color development. A solid yellow color was recorded as a positive reaction, and no color change was interpreted as a negative reaction. The ELISA plate was then analyzed by reading the absorbance using the Microplate read1000 (Global Diagnostics B, SA). Absorbance values that were two times the value of the negative control were recorded as positive reactions.

Figure 2.2: Map showing the areas in green circles selected as sampling sites around South Africa; the map was created using ArcGIS software

Table 2.1: Identity and locations of the samples used in this study.

Sample ID	Cultivar	Location	Coordinates
NW1	LT93	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW2	LT93	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW3	LT93	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW4	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW5	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW6	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW7	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW8	LT93	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW9	AUL280	Oudtshoorn	S33°35′ 0″
			E22°12′ 0″
WC10	AUL280	Oudtshoorn	S33°35′ 0″
			E22°12′ 0″
WC11	AUL280	Oudtshoorn	S33°35′ 0″
			E22°12′ 0″
LP12	LK26R	Groblersdal	S25°08'04.0”
			E29°18'23.5”
LP13	LK43/34	Groblersdal	S25°10’35.93”
			E29°23.35.87”
LP14	LD2	Ellisras	S23°12'0”
			E27°46'60”
LP15	LD2	Ellisras	S23°12'0”
			E27°46'60”
LP16	LD2	Ellisras	S23°12'0”

			E27°46'60”
LP17	LD2	Ellisras	S23°12'0”
			E27°46'60”
LP18	LD2	Ellisras	S23°12'0”
			E27°46'60”
LP19	LD2	Ellisras	S23°12'0”
			E27°46'60”
NW20	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW21	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”
NW22	BS4	Rustenburg	S25°43’28.3”
			E27°17’28.5”

2.2.4. RNA Extraction

Total RNA was extracted from the frozen symptomatic leaf material from both original field samples and inoculated *Nicotiana tabacum cv.* Xanthi and *N. benthamiana* plants using a Quick-RNA™ Plant Miniprep kit (ZYMO Research, USA) following the manufacturer's instructions. A volume of 30 µl of the total RNA was eluted and quantified using the Nanodrop 1000™ spectrophotometer (ThermoFisher Scientific™, USA) and kept at -80℃ until it was used for further analysis.

2.2.5. Reverse transcriptase-polymerase reaction

A volume of 5µl of total extracted total RNA was incubated at 65℃ in a Thermocycler for 5 minutes. The cDNA strand was synthesized using the RevertAid Reverse Transcriptase (RT) kit (ThermoFisher Scientific™, USA) and the 2mM of reverse primer in 20µL reaction (Table 2). The reverse transcription reactions were incubated at 42℃ for 60 minutes and 70℃ for 10 minutes in a Thermal cycler. Each PCR reaction was prepared by adding 10µl Dreamtaq 2x master mix, 4µl cDNA, and 1mM for each primer (forward and reverse; Table 2.2.) and filled up to a volume of 20 µL with nuclease-free water. Thirty-five PCR cycles were run as follows; 95℃ denaturation for 30 s, specific annealing temperature (Table 2.2) for each primer set 45 s, 72℃ elongation for 60 s, and a final extension for 10 min. The samples were tested using generic orthotospovirus primers (gM410 and gM870c) and TSWV-specific primers

(TSWV722 and TSWV723) (Table 2.2). The PCR products were resolved by agarose gel electrophoresis on a 1.5% gel pre-stained with SYBR safe DNA stain (Invitrogen, USA) run at 100 volts for 45 minutes. The gel was viewed under UV light using a G-Box image analyzer machine (Syngene, Cambridge, UK), and pictures were taken using Genesnap software.

2.2.6. Sanger Sequencing and phylogenetic analysis

The amplified PCR products were purified using the ZYMO Gel DNA Recovery kit as per manufactures' instructions. Purified PCR products of eight samples amplified with TSWV-specific primers and three amplified with generic orthotospovirus primers were sent for Sanger sequencing at KwaZulu-Natal Research and Innovation Sequencing Platform (KRISP) in Durban. The obtained sequences were subsequently subjected to the Basic Local Alignment Search Tool (BLASTn), one of the more popular choices for searching and aligning sequences (http://blast.ncbi.nlm.nih.gov/). Homologous sequences of virus isolates from the different SA provinces (obtained from this study), and sequences of previously identified isolates obtained from GenBank (Table 2.4) were selected and aligned in Clusta*W* for phylogenetic analyses based on the best fit model proposed by MEGA X software.

The evolutionary history was inferred using the Maximum Likelihood method and Tamura 3-parameter model (Tamura, 1992). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with a superior log-likelihood value. Evolutionary analyses were conducted in MEGA X (Kumar *et al.,* 2018).

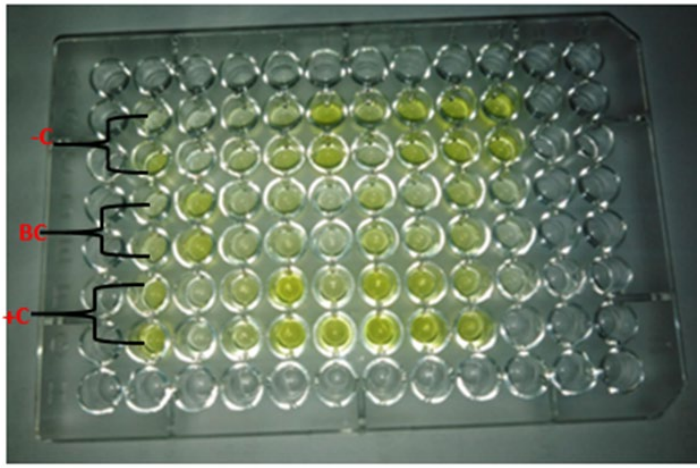
Table 2.2: Primers used in RT-PCR to detect orthotospoviruses and TSWV in tobacco samples from different provinces of SA.

Primer	5'-3'sequence	Annealing T°	Expected band size	reference
gM410 (forward)	AACTGGAAAAATGATTYNYTTGTTGG	52℃	500bp	Adkins and Rosskopf 2002.
gM870c (reverse)	ATTAGYTTGCAKGCTTCAATNAARGC			
TSWV722 (forward)	GCTGGAGCTAAGTATAGCAGC	55℃	620bp	Chen *et al*. 2012.
TSWV723 (reverse)	CACAAGGCAAAGACCTTGAG			

2.3. RESULTS

2.3.1. DAS-ELISA

All the 22 samples collected from the field were tested using orthotospovirus coat protein-specific antibodies, and 19 of those samples tested positive for orthotospoviruses (Table 2.3). Samples were considered positive if they had a solid yellow color (Fig. 2.3) or absorbance values 2X the average of the negative sample control. The positive samples had absorbance values above 0.2, while the buffer and negative control showed the absorbance value of 0.05 in the microplate reader.



.

Figure 2.3: ELISA plate showing positive and negative results. Positive results are identified by a distinct solid yellow color similar to that of the positive control (+C). A negative reaction is observed by faint yellow color similar to that of the buffer (BC) and negative control (-C).

Table 2.3: ELISA results of TSWV in the collected samples from 3 major tobacco-producing regions in South Africa

Sample	Cultivar	ELISA outcome	Locations
NW1	LT93	+	Rustenburg
NW2	LT93	+	Rustenburg
NW3	LT93	+	Rustenburg
NW4	BS4	+	Rustenburg
NW5	BS4	+	Rustenburg
NW6	BS4	+	Rustenburg

NW7	BS4	+	Rustenburg
NW8	LT93	+	Rustenburg
WC9	AUL280	+	Oudtshoorn
WC10	AUL280	-	Oudtshoorn
WC11	AUL280	+	Oudtshoorn
LP12	LK26R	+	Groblersdal
LP13	LK43/34	+	Groblersdal
LP14	LD2	+	Ellisras
LP15	LD2	+	Ellisras
LP16	LD2	+	Ellisras
LP17	LD2	+	Ellisras
LP18	LD2	+	Ellisras
LP19	LD2	+	Ellisras
NW20	BS4	-	Rustenburg
NW21	BS4	+	Rustenburg
NW22	BS4	-	Rustenburg

(+) indicates the samples that tested positive; (-) shows the samples that tested negative for TSWV

2.3.2. Biological assays

Forty-four plants were inoculated with the field samples (two plants for each sample). Sixteen showed typical virus symptoms: vein clearing, mosaic, leaf chlorosis, and necrotic lesions on stems and leaves. The symptom development was first observed seven days post-inoculation (dpi), and the first symptoms were vein clearing from samples 1,3,11,12,13,15, and 16. Eight dpi samples 15 and 16 also developed the vein-clearing symptoms, which later developed into severe mosaic symptoms (Fig. 2.4C). At nine dpi, samples 5 and 6 showed necrotic spots that developed to severe necrosis and stunted growth (Fig. 2.4B). The plants were monitored until they reached the flowering stage.

Figure 2.4: Mechanical inoculated *Nicotiana tabacum* plants ten days post-inoculation showed Chlorotic spots (A), Necrotic lesions with leaf curling (B and D), and mosaic symptoms (C).

2.3.3. RT-PCR

Nineteen of the 22 symptomatic field leaf samples tested using the orthotospovirus primers produced the expected band size of ~500bp (Fig 2.5). Eleven of the 22 samples tested using TSWV specific primers produced an expected band size of ~620bp (Fig. 2.6). None of the mechanically inoculated plants tested positive for orthotospoviruses when using the Orthotospovirus generic primers.

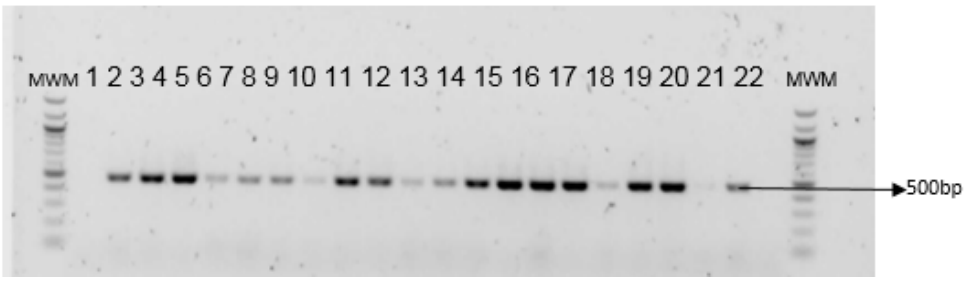


Figure 2.5: Agarose gel showing the RT-PCR products from all the samples tested with orthotospovirus generic primers (gM410 and gM870c); and the molecular weight marker (on the far left and right ends of the gel picture) There were 19 of the tested samples that showed the expected bands of approximately 500bp

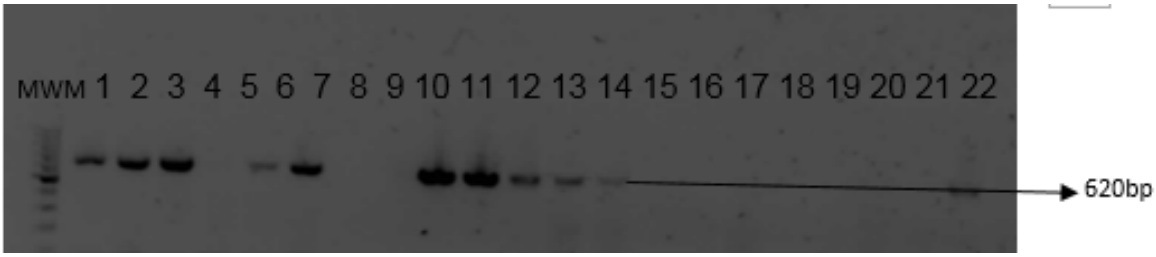


Figure 2.6: Agarose gel showing the RT-PCR products from all the samples tested with TSWV722 & TSWV723 primers; the molecular weight marker (on the far-left end of the gel picture). 11 samples of the tested that showed the expected band sizes of approximately 620bp

2.3.4. Sequencing

Eight amplicons were obtained using TSWV specific primers, and three amplicons obtained using generic orthotospovirus primers were validated by Sanger sequencing. The obtained sequences were subjected to BLASTn, which showed that the three amplicons; [LP14 (accession no.: OL505552), LP18 (accession no.: OL505553), and NW4 (accession no.: OL505554)] obtained using generic orthotospovirus primers were closely related to GRSV NSm gene with nucleotide identities of 96-97% (Table 2.4). The isolates found in this study; NW4 (accession no: OL471727), NW6 (accession no: OL471728), NW2 (accession no: OL4717279), NW1 (accession no: OL471730), NW3 (accession no: OL471734) NW21 (accession no: OL471729), WC9 (accession no: OL471732) and WC11 (accession no: OL471733) matched 98% with the nucleocapsid gene (N-gene) of Zimbabwean and South African TSWV isolates (Table 2.4).

Table 2.4: Blast analysis results of the TSWV N-gene and GRSV NSm-gene partial sequences detected in this study

Isolates (this study)	Nucleotide Identity match %	Closest isolate	Origin	Accession no.	Reference
NW4	98.19 to TSWV N-gene	Mum-A5	Zimbabwe	MG602673	Karavina *et al.,* 2020
NW6	97.91 to TSWV N-gene	Mum-A5	Zimbabwe	MG602673	Karavina *et al.,* 2020
NW2	98.05 to TSWV N-gene	Mum-A5	Zimbabwe	MG602673	Karavina *et al.,* 2020
NW1	98.51 to TSWV N-gene	TSWV-KZN	SA	DQ834847	Siviparasad & Gubba 2008)
NW21	98.44 to TSWV N-gene	TSWV-KZN	SA	DQ834847	Siviparasad & Gubba 2008)
NW3	98.02 to TSWV N-gene	77-pep-zim	SA	KU671049	Unpublished
WC11	98.15 to TSWV N-gene	98/0472	SA	AJ296600	Heinze *et al,* 2001
WC9	98.77 to TSWV N-gene	77-pep-zim	Zimbabwe	KU884649	Karavina *et al.,* 2020
LP14	97.25 to GRSV NSm-gene	SA-05	SA	MH742957	Silva *et al.,* 2019
LP18	96.20 to GRSV NSm-gene	SA-05	SA	MH742957	Silva *et al.,* 2019
NW4[a]	96.72 to GRSV NSm-gene	SA-05	SA	MH742957	Silva *et al.,* 2019

2.3.5. Phylogeny

The phylogenetic analysis of representative sequences of isolates generated in this study was conducted using 23 closely related TSWV and 10 GRSV related sequences (Table 2.5). The cucumber mosaic virus was used as outgroup to root the trees. Phylogenetic analysis for the N

gene of the different TSWV isolates resulted in the generation of two major groups: the South African, European, and Asian group on the one hand and the Zimbabwean group on the other (Fig 2.7). The isolates generated in this study (NW21, WC11, NW3, and WC9 all clustered with the South African, European, and Asian groups. Within this group; isolates NW3 and WC9 formed a distinct clade with a bootstrap value of 96%; isolate WC11 clustered with isolate 98/0472 with a bootstrap value of 91%, and isolate NW21 that was closely related to a cluster of previously identified South African TSWV isolates; TSWV-KZN, TSWV-NW1, and TSWV-NW2, with a bootstrap value of 69%. Phylogenetic analysis of the NSm gene of GRSV showed that the isolates from this study (LP14 and NW4) formed a distinct clade with the bootstrap value of 96%. These isolates were related to a cluster that included the South African GRSV isolate SA-05 (MH742957) with a bootstrap value of 75% (Fig. 2.8).

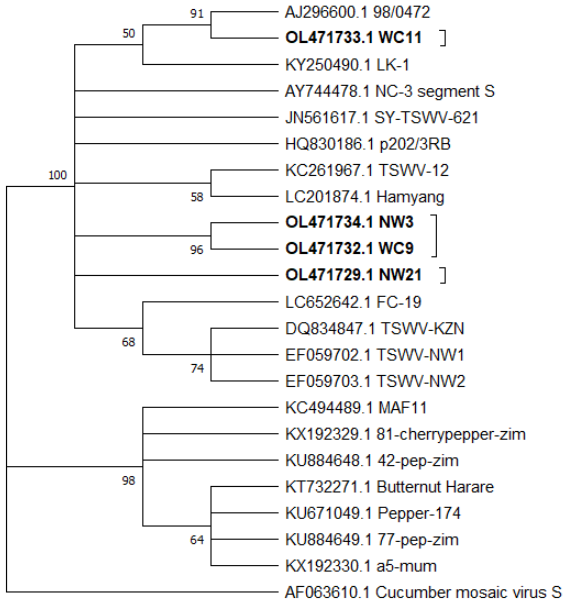


Figure 2.7: Phylogenetic analysis of the partial N-gene of the TSWV isolates (NW21, WC11, NW3, and WC9) infecting tobacco in South Africa and the isolates available in the NCBI database from other geographical regions with the cucumber mosaic virus (CMV) sequence as an outgroup. (SA isolates from this study are printed in bold font). The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

Table 2.5. Isolates obtained from the NCBI database of TSWV, GRSV, TCSV, and CMV nucleotide sequences were used for phylogenetic analysis in conjunction with isolates from this study

Virus	Isolate	Origin	Accession no.	Rererence
TSWV	WC9	SA	OL471732	This study

TSWV	NW21	SA	OL471729	This study
TSWV	NW3	SA	OL471734	This study
TSWV	WC11	SA	OL471733	This study
TSWV	TSWV-NW1	SA	EF059702	Siviparasad & Gubba 2008
TSWV	98/9472	SA	AJ296600	Heinze *et al .,* 2001
TSWV	TSWV-NW2	SA	EF059703	Sivparasad & Gubba 2008
TSWV	LK-1	SA	KY250490	Unpublished
TSWV	a5-mum	Zimbabwe	KX192330	Unpublished
TSWV	Pepper-174	Zimbabwe	KU671049	Unpublished
TSWV	Buttenut harare	Zimbabwe	KT73271	Unpublished
TSWV	gyp-316-zim	Zimbabwe	KX273062	Unpublished
TSWV	92-tom1-zim	Zimbabwe	KU892656	Unpublished
TSWV	Cuc-48-zim	Zimbabwe	KX273061	Unpublished
TSWV	86-potato-zim	Zimbabwe	KX192332	Unpublished
TSWV	90-eggplant-zim	Zimbabwe	KX192331	Unpublished
TSWV	81 cherrypepper	Zimbabwe	KX192329	Unpublished
	77-pep-zim	Zimbabwe	KU884649	Unpublished
TSWV	42-pep-zim	Zimbabwe	KU884681	Unpublished
TSWV	Pepper-174	Zimbabwe	KU671049	Unpublished
TSWV	LYE89	France	FR693062	Tentchev *et al,* 2011
TSWV	MAF11	New Zealand	KC494489	Timmerman-Vaughan *et al.,* 2014
TSWV	TSWV12	South Korea	KC261967	Lian *et al.,* 2013
TSWV	Hamyang	South Korea	LC201874	Unpublished
TSWV	p202/3RB	Italy	HQ830186	Unpublished
TSWV	SY-TSWV-621	Syria	JN561617	Unpublished
TSWV	H2	Poland	KU308369	Unpublished
TSWV	TSWV-TK1	Poland	KR186203	Unpublished
TSWV	LSJo1	Jordan	JN601519	Unpublished
TSWV	Is-334	Montenegro	GU369716	Unpublished
GRSV	BR	Brazil	KY35037	Unpublished
GRSV	LP14	SA	OL505552	This study
GRSV	NW4	SA	OL505553	This study
GRSV	LEM	Brazil	MH686227	De Marchi *et al.,* 2019
GRSV+TCSV	LGMTSG**	USA	HQ644141	Webster et al., 2011
GRSV	S4	Brazil	MH388796	Unpublished
GRSV	S30	Brazil	MH388798	Unpublished
GRSV	-	Brazil	AF513220	Unpublished
GRSV	GRSV-BR-PI	Brazil	MK503850	Unpublished
GRSV	S8	Brazil	MH388797	Unpublished
GRSV	SA-05	SA	MH742957	Silva et al.,2019
GRSV	GRSV-AR	Argentina	KT972591	de Breuil *et al.,* 2015
CMV	S	SA	AF063610	Unpublished

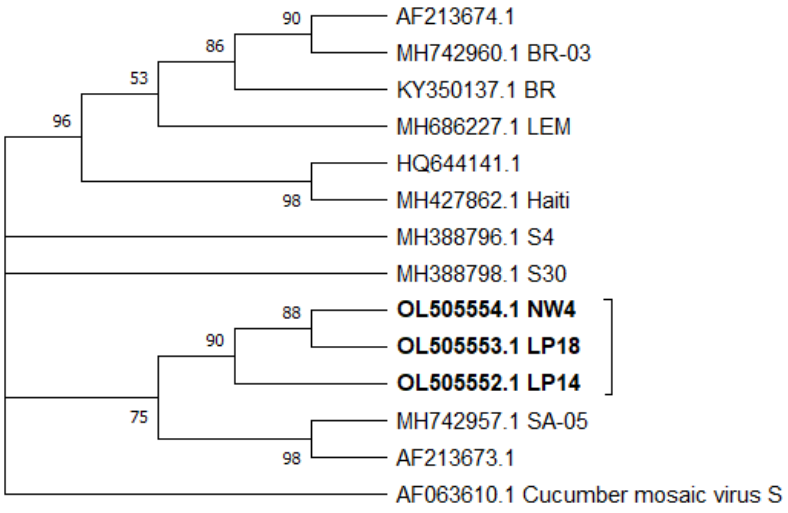


Figure 2.8: Phylogenetic relationships of partial Nsm gene sequence of the GRSV isolates infecting tobacco in SA (NW4, LP18, and LP14)and the isolates available in the NCBI database from other geographical regions with the CMV sequence as an outgroup .isolates from this study are printed in bold font. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree inferred from 1000 replicates

2.4. DISCUSSION

Orthotospoviruses are considered some of the most destructive groups of plant pathogens worldwide (Oliver & Whitfield, 2016), with TSWV alone in the top ten economically important plant viruses causing losses of over 1 billion US dollars for several crops worldwide (Rybicki,2015). The results presented in this study confirmed the presence of Orthotospoviruses in tobacco. They were detected with the orthotospovirus-specific coat protein antibodies (Table 2.3). The ELISA results were in agreement with the RT-PCR results (Figure 2.5 & 2.6). The PCR products were further characterized using Sanger sequencing and phylogenetic analysis (Table 2.4; Fig 2.7& 2.8), which then identified the two orthotospoviruses; TSWV and GRSV., This study, along with other studies that have been done on orthotospoviruses in SA (Moore, 1933; Kisten *et al.,* 2016; Silva *et al.,* 2019; Sivparsad and Gubba, 2008), confirms that the presence of orthotospoviruses which pose a significant threat to tobacco production in South Africa.

GRSV was detected in samples from Limpopo and Northwest in this study. To our knowledge, this is the first report of GRSV on tobacco in SA since it was first reported in *Arachis* species in 1966 and soybean in 2002 (Klesser, 1966; Petersen *et al.,* 2002). In general, phylogenetic

analysis showed that the TSWV and GRSV isolates generated from this study were related to previously identified South African isolates of both viruses (Figs. 2.7 and 2.8). This finding probably points to the fact that there has been very little if any mutation of both viruses in South Africa over the years. This has implications on how diseases caused by the two viruses on different crops can be managed.

The bioassays were done to propagate the viruses, so there will be enough plant material to conduct the study. However, not all the inoculated plants developed symptoms. Some of the symptoms on the propagation hosts (*Nicotiana tabacum* cv. Xanthi nc. and *Nicotiana benthamiana*) were typical of those produced by viruses; the symptoms included mosaic, interveinal necrosis, chlorotic spots, and severe necrosis, which led to death in some plants (Fig 2.3A-D). When the inoculated leaf material was tested with RT-PCR, and they were all negative for Orthotospoviruses. This can be explained by the fact that the Orthotospoviruses in the field samples lost their infectivity when they were stored before use. Orthotospoviruses are known for their instability, and it thus becomes harder to mechanically transmit virus particles (Halliwell and Philley, 1974; Mandal *et al.,* 2001). The symptoms observed on the inoculated plants could have been caused by more stable viruses in the mixture of viruses infecting the tobacco plants.

In conclusion, the methods used in this study confirm the presence of two orthotospoviruses in the tested field samples, with TSWV being present in all the provinces and GRSV present in Limpopo and Northwest. This is the first time GRSV has been reported on tobacco in SA. It should be pointed out that, under field conditions, mixed infections are very common. Several factors contribute to the emergence of mixed plant virus infections, including genetic changes, recombinations or reassortments in the virus genome, insect vector population changes, and migration due to crop trade and climate conditions changes (Dennehy, 2017). The methods used to identify the virus infecting tobacco might be telling only part of the story. To that end, more effective and more accurate identification methods such as Next-generation sequencing (NGS) must be employed to obtain the complete picture of the viruses infecting a crop at a particular time. Only then can efficient and sustainable control virus disease strategies be developed.

2.5. REFERENCES

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

Chen, T.C., Li, J.T., Lin, Y.P., Yeh, Y.C., Kang, Y.C., Huang, L.H. and Yeh, S.D., 2012. Genomic characterization of Calla lily chlorotic spot virus and design of broad-spectrum primers for detection of tospoviruses. *Plant Pathology*, *61*(1): 183-194.

Culbreath, A.K., Todd, J.W., and Brown, S.L., 2003. Epidemiology and management of tomato spotted wilt in peanut. *Annual Review of Phytopathology* 41: 53–75.

de Breuil, S., La Rossa, F.R., Giudici, A., Wulff, A., Bejerman, N., Giolitti, F. and Lenardon, S., 2015. Phylogenetic analysis of Groundnut ringspot virus isolates from peanut and identification of potential thrips vectors in peanut crop in Argentina. *AgriScientia*, *32*(1): 77-82.

de Haan P., Wagemakers L., Peters D. and Goldbach R., 1990. The S RNA segment of the tomato spotted wilt virus has an ambisense character. *Journal of General Virology* 71: 1001-1007.

De Marchi, B.R., Bello, V.H., Watanabe, L.F.M., da Silva, F.B., Müller, C., Pavan, M.A. and Krause-Sakate, R., 2019. Characterization and complete genome sequence of groundnut ringspot Orthotospovirus in soybean in Brazil. *Journal of Plant Pathology*, *101*(2): 401-401.

Dennehy, J.J., 2017. Evolutionary ecology of virus emergence. *Annals of the New York Academy of Sciences*, *1389*(1): 124.

Dong, J.H., Yin, Y.Y., Xu, X.Y., Duan, Y.M. and Zhang, Z.K., 2010. First report of Tomato spotted wilt virus in tomato and tobacco in China. *Journal of Plant Pathology*, *92*(4): 121

EPPO 2020: https://gd.eppo.int/taxon/1TOSPG accessed 07 September 2021

FAOSTAT <http://www.fao.org/faostat/en/#data/QV accessed 26 September 2021

Grisan, S., Polizzotto, R., Raiola, P., Cristiani, S., Ventura, F., Di Lucia, F., Zuin, M., Tommasini, S., Morbidelli, R., Damiani, F. and Pupilli, F., 2016. Alternative use of tobacco as a sustainable crop for seed oil, biofuel, and biomass. *Agronomy for Sustainable Development*, *36*(4): 1-8.

Halliwell, R.S. and Philley, G., 1974. Spotted wilt of peanut in Texas. *Plant Disease Reporter*, *58*(1): 23-25.

Heinze, C., Letschert, B., Hristova, D., Yankulova, M., Kauadjouor, O., Willingmann, P., Atanassov, A. and Adam, G., 2001. Variability of the N-protein and the intergenic region of

the S RNA of tomato spotted wilt tospovirus (TSWV). *The New Microbiologica*, *24*(2): 175-187.

ICTV online: https://talk.ictvonline.org/taxonomy/ accessed 05 May 2021

Karavina, C., Ibaba, J.D. and Gubba, A., 2020. Full genome sequence of a chrysanthemum-infecting tomato spotted wilt virus isolate from Zimbabwe obtained by next-generation sequencing. *Acta Virologica*, *64*(1): 88-92.

Kaye, A.C., Moyer, J.W., Parks, E.J., Carbone, I. and Cubeta, M.A., 2011. Population genetic analysis of Tomato spotted wilt virus on peanut in North Carolina and Virginia. *Phytopathology*, *101*(1): 147-153.

Klesser, P.J., 1966. Groundnut ringspot virus-a new sap-transmissible virus of Arachis hypogaea and A monticola. *South African Journal of Agricultural Science*, *9*(3): 711-720.

Kormelink, R., van Poelwijk, F., Peters D., and Goldbach, R., 1992. Non-viral heterogeneous sequences at the 5' ends of tomato spotted wilt virus mRNAs. *Journal of General Virology* 73: 2125-2128

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6): 1547-1549.

Lian, S., Lee, J.S., Cho, W.K., Yu, J., Kim, M.K., Choi, H.S. and Kim, K.H., 2013. Phylogenetic and recombination analysis of tomato spotted wilt virus. *PloS one*, *8*(5): 63380.

Mandal, B., Pappu, H.R. and Culbreath, A.K., 2001. Factors affecting mechanical transmission of Tomato spotted wilt virus to peanut (Arachis hypogaea). *Plant Disease*, *85*(12): 1259-1263.

Moore, E.S. and Andessen, E.E., 1939. Notes on Plant Virus Diseases in South Africa. I. The Kromnek Disease of Tobacco and Tomato. Dept. Agric., Union of South Africa. *Scientific. Bull*, *182*: 1-36.

Oliver, J.E. and Whitfield, A.E., 2016. The genus Tospovirus: emerging bunyaviruses that threaten food security. *Annual Review of Virology*, *3*: 101-124.

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.

Pietersen, G. and Morris, J., 2002. Natural occurrence of Groundnut ringspot virus on soybean in South Africa. *Plant Disease*, *86*(11):1271-1271.

Silva, J.M.F., de Oliveira, A.S., de Almeida, M.M.S., Kormelink, R., Nagata, T. and Resende, R.O., 2019. Tomato chlorotic spot virus (TCSV) putatively incorporated a genomic segment of groundnut ringspot virus (GRSV) upon a reassortment event. *Viruses*, 11(2): 187.

Sivparsad, B.J. and Gubba, A., 2008. Isolation and molecular characterization of Tomato spotted wilt virus (TSWV) isolates occurring in South Africa. *African Journal of Agricultural Research*, *3*(6): 428-434.

Status of Tobacco Production and Trade in Africa – Factsheets (August 2020) NTM code Source: Consolidated Tariff Schedules database (CTS), http:\\tdf.wto.org Source: Most recent national survey available to WHO: 2019; STEPS Survey 375 16 % of agricultural land Source: http://wits.worldbank.org/ accessed 06 August 2021

Tamura, K., 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. *Molecular Biololgy Evolution*, *9(4):* 678-687.

Tentchev, D., Verdin, E., Marchal, C., Jacquet, M., Aguilar, J.M. and Moury, B., 2011. Evolution and structure of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes. *Journal of General Virology*, *92*: 961-973.

Timmerman-Vaughan, G.M., Lister, R., Cooper, R. and Tang, J., 2014. Phylogenetic analysis of New Zealand tomato spotted wilt virus isolates suggests likely incursion history scenarios and mechanisms for population evolution. *Archives of Virology*, *159*(5): 993-1003.

Turina, M., Kormelink, R. and Resende, R.O., 2016. Resistance to tospoviruses in vegetable crops: epidemiological and molecular aspects. *Annual Review of Phytopathology*, *54*: 347-371.

Webster, C.G., Reitz, S.R., Perry, K.L. and Adkins, S., 2011. A natural M RNA reassortant arising from two species of plant-and insect-infecting bunyaviruses and comparison of its sequence and biological properties to parental species. *Virology*, *413*(2): 216-225.

Whitfield, A.E., Ullman, D.E. and German, T.L., 2005. Tospovirus-thrips interactions. *Annual Review of Phytopathology* ., *43*: 459-489.

Chapter 3: Identification of viruses and genome analysis of orthotospoviruses occurring in mixed infections on tobacco (*Nicotiana* sp.) in South Africa

ABSTRACT

Viruses are among the most destructive plant pathogens worldwide. Therefore, accurate and efficient detection methods and genome characterization are crucial for developing effective virus disease control strategies. This study aimed to identify viruses occurring in mixed infections on tobacco and fully characterize the three orthotospoviruses identified using Next Generation Sequencing (NGS). The obtained (NGS) data were analyzed using the Genome Detective Virus Tool. The assembled contigs were subjected to Blastn, and those matching with orthotospovirus sequences were selected for further analysis. The NGS results showed 15 viruses infecting tobacco in the samples analyzed; nine of those viruses are first reports in SA. The whole TSWV genome was obtained and was analyzed using Clustal*W*. The TSWV L segment for TSWV_Nto isolate (from this study) was closely related to LK-1 SA isolate; the M segment was closely related to DSMZ PV-0393 isolate from China, and the S segment was closely related to the p20/3WT isolate from Italy. All the TCSV and GRSV segments were closely related to those of other SA isolates previously identified; DSMZ PV-0205 and SA-05. To our knowledge, this is the first report of TCSV and GRSV infecting tobacco in SA. Recombination analysis was done using RDP4. Various recombination events were discovered among some of the isolates generated from this study, indicating that two or more different isolates of one virus infect the same host. In conclusion, the results from this study show that it is important to use technologies such as NGS to have an accurate picture of the whole virus population infecting a specific crop for proper disease monitoring strategies and the development of efficient control strategies

3.1. INTRODUCTION

Plant viruses and viroids are amongst the most important plant pathogens that cause a significant decrease in crop yield quality and economic losses worldwide (Zhao *et al.,* 2020). Most of those losses are due to emerging and recurrent viruses. Complex factors associated with viral disease outbreaks, including rapid evolution, insect vector migration, and unpredictable increase of viral host range, make it very difficult to develop effective long-term disease management strategies (Zaidi *et al.,* 2016). Therefore, early and effective identification methods remain the focal point in virology and help combat and reduce the negative impact on crop production (Flint *et al.,* 2020).

Orthotospoviruses are among the most damaging plant viruses causing severe necrotic lesions and wilting symptoms on several crops (Pappu *et al.,* 2009). Tomato spotted wilt orthotospovirus (TSWV), groundnut ringspot orthotospovirus (GRSV), and tomato chlorotic spot orthotospovirus (TCSV) have been identified as the most prevalent orthotospoviruses since the early 1990s (Martínez *et al.,* 2019). The viruses share a similar host range and have the same thrip vector species for efficient transmission (Resende *et al.,* 2004). Orthotospovirus economic importance is due to their damaging ability, and that they are difficult to control due to their wide host range and mode of transmission (Martínez *et al.,* 2018).

Orthotospoviruses have a tripartite negative-sense RNA genome which consists of three RNA segments; Small (S), Medium (M), and Large (L). They use an ambisense replication strategy to produce proteins, coding for five proteins. The L RNA encodes for the RNA-dependent RNA polymerase in a negative sense (de Haan *et al.,* 1990; Kormelink *et al.,* 1992). The M RNA encodes for two proteins: the non-structural movement protein (NSm) in the positive direction and Gn-Gc glycoprotein in the negative direction (de Haan *et al.,* 1990). The S RNA encodes for the non-structural protein (NSs) in the positive sense and the nucleocapsid (N) protein in the negative sense (Whitfield *et al*., 2005; Pappu *et al*., 2009). Two studies have been done on the whole genome of TSWV (isolate LK1) and GRSV (SA-05) in SA (Kisten, 2016; Silva *et al.,* 2019).

Reliable plant virus detection methods are important for improved crop protection; there are specific and non-specific methods. Specific methods usually require prior knowledge of the pathogen being tested, while non-specific methods do not require any specific knowledge of the pathogens (Adams *et al.,* 2016). The adoption of molecular techniques such as reverse transcription-polymerase chain reaction (RT-PCR) has increased the speed and accuracy of

viral disease diagnosis in crops. However, these techniques only detect known viruses, and each test is limited to one or a small number of related viruses (Massart *et al.,* 2014).

The use of new technologies for virus identification has increased the understanding of viral genome variability and evolution over the years (Barba *et al*., 2014). Unlike the traditional detection techniques, e.g., enzyme-linked immunosorbent assay, Polymerase chain reaction, or bioassays, the use of Next-generation sequencing (NGS) technology efficiently and accurately determines the causative pathogen without prior knowledge of the disease pathogen (Villamor *et al.,* 2020). The plant virus genome can be detected rapidly even when in very low concentrations in an infected host. NGS can detect both DNA and RNA viruses, and it also makes it possible to assemble partial or complete virus sequences. NGS is one of the most accurate and time-efficient detection methods for virus identification (Akinyemi *et al.,* 2016).

In cases where mixed virus infections occur, it becomes important to identify the individual viruses that are part of the mixture if effective control strategies to manage the ensuing viral diseases are to be developed. NGS presents great advantages in investigating the molecular interactions of viruses with their hosts through transcriptome sequencing (RNA-seq). This study established (Chapter 2) that more than one virus contributed to the symptoms observed on tobacco in the field. Against this background, this study aimed to: 1) identify viruses infecting tobacco in SA using NGS, 2) analyze the genomes of the identified orthotospoviruses, and 3) study the relationships between the identified orthotospoviruses and other known orthotospoviruses.

3.2. MATERIALS AND METHODS

3.2.1. Collection and sample prep

Samples were collected and prepared as described in Chapter 2.

3.2.2. RNA extraction and sequencing

Total RNA was extracted from the symptomatic leaf material using a Quick-RNA™ Plant Miniprep kit (ZYMO Research, USA), following the manufacturer's instructions. A volume of 50 µl of the total RNA was eluted, quality checked, and the concentration was measured using a Nanodrop 1000™ spectrophotometer (ThermoFisher Scientific™, USA). To cut costs, the RNA from 22 collected samples was combined into one Eppendorf tube and kept at -80℃ until it was sent in dry ice to Agricultural Research Council (Biotechnology Platform) for NGS library preparation and sequencing. The sample was run on an Illumina Hiseq2500 Ultra-High

Throughput Sequencing system (Illumina Inc. USA). The 151x151bp read pair-end library was generated using the platforms' sequencing by synthesis technology (SBS).

3.2.3. Sequence analysis

The generated NGS data quality was assessed using FastQC version 0.11.9. The pair end of the data was analyzed using the genotyping online pipeline software; Genome detective virus tool version 1.133 (Vilsker *et al.,* 2019) embedded with Trimmomatic (Bolger *et al.,* 2014), FASTQC (Brown *et al.,* 2017), DIAMOND (Buchfink *et al.,* 2015), SPAde (Bankevich *et al.,* 2012) and Advanced genome aligner (AGA) (Deforche, 2017). The tool enables the analysis of whole or partial viral genomes within minutes. Genome Detective is a web-based pipeline that allows raw NGS data to be assembled into de novo complete viral genomes quickly and accurately. Before the *de novo* assembly, the reads that were not viral related were filtered out. The application uses a novel alignment method, AGA that constructs genomes by reference-based linking *de novo* contigs by combining amino acids and nucleotide scores. The contigs that matched Orthotospoviruses genomes were selected and aligned using Clustal*W* found in Molecular Evolutionary Genetics Analysis (MEGA X, Kumar, *et al.,* 2018) software to generate consensus sequences for the L, M, and S segments. The open reading frames (ORFs) for each sequence were determined with the OFR finder in the NCBI. The molecular weight of proteins was determined using the ExPASy bioinformatics tool (https://web.expasy.org/protparam/) (Geisteiger *et al.,* 2005).

3.2.4. Phylogeny

The consensus sequences were subsequently subjected to the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/), one of the more popular choices for searching and aligning sequences. The sequences of isolates generated in this study and isolates from different countries obtained from the NCBI database were selected and aligned for phylogenetic analyses based on the best fit model proposed by MEGA X software (Kumar *et al.,* 2018). The evolutionary history was inferred using the Maximum Likelihood method and Tamura-Nei model (Tamura and Nei, 1993). The bootstrap consensus tree was inferred from 1000 replicates taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). Initial tree(s) for the heuristic search were automatically obtained by applying Neighbor-Join and

BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model and selecting the topology with a superior log-likelihood value.

3.2.5. Recombinant analysis

The recombinant events for each virus were tested using the Recombinant Detection Program version4.101 (RDP4) embedded with RDP (Martin and Rybicki, 2000), BootScan (Martin *et al.,* 2005), MaxChi (Smith, 1992), SiScan (Gibbs *et al.,* 2000), Chimaera (Posada and Crandall, 2001), GENECONV (Padidam *et al.,* 1999) and 3Seq (Boni *et al.,* 2007). Segmental recombination events of the isolates were analyzed based on at least four positive statistical methods embedded in the RDP4 program. The nucleotide sequences that were used for phylogenetic analysis were also used for recombinant analysis for each virus.

3.3. RESULTS

3.3.1. Quality control and analysis

The total RNA concentration of the sample was 738.1 ng/µl, and the absorbance ratio at A260/A280 was 2.00, which is marginally equal to the standard value of 2.0 when the RNA sample is pure. The NGS data generated was 4.28 gigabytes, which consisted of 47 106 512 raw reads (Table 3.1). The filtering started with 42 939 825 reads, 40 910 808 reads (95%) that did not appear to be viral were removed. The *de novo* assembly started with 2 029 050 (5% of the total data) reads (Table 3.1); about 64% were assembled into viral contigs.

Table 3.1: Characteristics of the Next-Generation Sequencing data generated from total RNA extracted from tobacco sample

Statistics	**Sample characteristics**
Total raw reads	47 106 512
Average length	151
Number of reads after trimming	42 939 825
Number of reads after host removal	2 029 050
Number of reads mapped to viral contigs	129 3105
Average length after trimming	136
Total viral contigs	488
Contigs matched plant viruses	105

3.3.2. Virus identification

Fifteen plant viruses and one virus satellite were identified using NGS (Table 3.2). NGS also was able to pick partial and whole sequences of these viruses with coverage, nucleotide identity with the reference sequence, or amino acid identity of between 70-100%. The viruses identified were; tobacco mosaic virus(TMV), Western African asystasia virus 1 (WAAV1), potato virus Y (PVY), tobacco vein clearing virus (TVCV), tomato leaf curl Uganda virus (ToLCUV), petunia vein clearing virus (PVCV), cucumber mosaic virus (CMV), groundnut ringspot virus (GRSV), tomato spotted wilt virus (TSWV), tomato chlorotic spot virus (TCSV), beet mosaic virus (BMV), beet western yellows virus (BWYV), beet cryptic virus 2 (BCV-2), beet cryptic virus 3 (BCV-3), and Cotton leaf curl Multan satellite virus (CLCuMuB). The analysis was able to recover whole genome sequences for TMV, PVY, CMV, and TSWV.

3.3.3. Sequencing

These virus nucleotide sequences identified were subjected to BLASTn to find their closely related isolates based on nucleotide identity percentages Genebank (Table 3.3). All the identified virus sequences were subjected to BLASTn to determine the sequences they are closely related to, based on nucleotide (nt) identity. The BLASTn results showed that for all the identified viruses from this study, there was an 80-99% nucleotide identity with their closely related isolates except for PVCV-SA which had 73.75% WAAV1_Bon had 77% (Table 3.3).

3.3.4. Genome organization

The data analysis showed that the three orthotospovirus RNA segments (L; M; and S) code for five proteins (Tables 3.4 &3.5). The genome detective online pipeline analysis identified these proteins and determined their polarity and size based on the number of amino acids. The molecular weight of these proteins are presented (Tables 3.4 and 3.5) and was determined using ExPASy bioinformatic tool. The size of the L segments for both TSWV and GRSV was approximately 8.6kb. The size of the M segments was approximately 4.8kb for both viruses. However, the S segments were slightly different in length for all three viruses; 2.7kb for TSWV, 3.3kb for TCSV

Table 3.2: Details of the partial and complete nucleotide sequences of viruses picked by NGS

Assignment	# Contigs	# Reads	Coverage (%)	Depth of Coverage	Sequence length (bp)
Tobacco mosaic virus	1	853802	99.9	17247.2	6386
Groundnut ringspot virus (segment L)	1	95583	100	1429.6	8625
West African Asystasia virus 1 (segment DNA B)	2	92766	91.9	4969.6	2446
Potato virus Y	1	38448	100	511.8	9696
Tobacco vein clearing virus	5	9976	71.4	228.9	5547
Tomato chlorotic spot virus (segment S)	1	9768	100	418.4	3309
Tomato leaf curl Uganda virus	2	9267	100	434.5	2720
Petunia vein clearing virus	2	7175	73.8	172.3	5317
Cucumber mosaic virus (segment RNA 3)	1	5736	98.7	338.2	2188
Tomato chlorotic spot virus (segment M)	1	4179	99.3	112.2	4814
Cucumber mosaic virus (segment RNA 2)	1	2973	99.2	126.9	3025
Tomato spotted wilt virus (segment L)	4	2390	97.4	35.5	8668
Cucumber mosaic virus (segment RNA 1)	1	1691	99.3	65.4	3335
Tomato spotted wilt virus (segment M)	1	1486	99.4	39.8	4793
Tomato spotted wilt virus (segment S)	2	1468	92.6	70.1	2699
Cotton leaf curl Multan virus satellite U36-1	2	1013	90.7	108.7	1225
Beet mosaic virus	8	301	77.3	5.2	7410
Beet western yellows virus	5	70	31.9	5	1807
Beet cryptic virus 2 (segment dsRNA2b)	2	27	53	4.3	807
Beet cryptic virus 2 (segment dsRNA2a)	2	23	42.4	4.4	674
Beet cryptic virus 3	1	2	13	1.2	207

Table 3.3: BLASTn analysis of the virus isolates detected by High-throughput/NGS sequencing analysis

Viruses identified	Isolate (this study)	Accession no. (this study)	% nt Identity match	Closest isolate	Accession no.	Origin	Reference
TMV	TMV_BS	OL471710	98.79	DSMZ OV-0943	MW854284	Germany	Unpublished
GRSV L segment	BonNDSA_GRSV	OK493952	97.19	SA-05	MH742956	South Africa	Silva *et al.,* 2019
PVY	Nto_PVY	OL471715	93.62	Q19-Harare	MG602676	Zimbabwe	Karavina *et al.,*2021
WAAV1	WAAV1_Bon	OL471720	77.10	Cameroon_wild cassava-29-14	KT444606	Cameroon	Leke *et al.,* 2016
ToLCUV	ToLCUV_SA	OL471710	81.03	FLATZ15A9Sph_C6	MN381113	Tanzania	Unpublished
TVCV	TVCV_sa	OL598362	83.83	taxon:107324	AF190123	USA	Lockhart *et al.,*2000
PVCV	PVCV_sa	OL598363	73.75	Kr	MK472692	South Korea	Unpublished
CMV RNA 1	CMV_BS	OL471711	98.80	Fny-CMV	D00356	New York	Rizzo and Palukaitis 1989
RNA 2	CMV_BS	OL471712	98.48	Fny-CMVRNAII2	D00355	New York	Rizzo and Palukaitis 1988
RNA 3	CMV_BS	OL471713	99.18	Fny	D10538	New York	Owen *et al.,* 1990
TCSV M	BonNDSA_TCSV	OK493953	96.09	DSMZ PV-0205	MT723992	South Africa	Silva *et al.,* 2019
S	BonNDSA_TCSV	OK493954	96.15	SA-05	MH742958	South Africa	Silva *et al.,* 2019
TSWV L	TSWV_Nto	OL471721	96.24	LK-1	KY250488	South Africa	UnphublishedUnpublished
M	TSWV_Nto	OL471722	96.85	DSMZ PV-0393	MW854272	China	Margaria *et al.,*2014
S	TSWV_Nto	OL471723	98.50	p202/3WT	HQ830187	Italy	Unpublished
BMV	BMVt_sa	OL598364	94.30	BtMV-Wa	AY206394	USA	Nemchinov *et al.,* 2004
BWYV	BWYV_sa	OL598365	88.39	Rouen 1	KU521325	France	Unpublished
BCV-2_sa2 dsRNA2a	BCV_sa2	OL598367	95.49	BCV-2-JKI 1825732	MK731958	Germany	Unpublished
BCV_sa2 dsRNA2b	BCV_sa2	OL598368	99.83	BCV2-JKI 1825732	MK731957	Germany	Unpublished
BCV_sa3	BCV_sa3	OL598369	99.04	Reference seq.	NC_043394	USA	Xie et al., 1993
CLCuMuB	CLCuMV_tobSA	OL598366	98.20	U36-1	AJ620463	UK	Unpublished

Table 3.4: The genome organization for the TSWV isolates found in tobacco (*Nicotiana tabacum*) in SA

Segment	Segment Length(nt)	ORF position(nt)	ORF length(nt)	Polarity	Protein coded	Protein MW (kDa)	No. of Amino acids
L	8668	33-8661	8628	-	RdRp	331.50	2876
M	4796	241-1009	768	+	Nsm	32.94	256
		4737-1330	3407	-	Gn-Gc	127.318	1136
S	2699	89-1483	1394	+	NSs	52.45	465
		2763-1987	776	-	N	28.84	259

Table 3.5: The genome organization for the GRSV L segment and TCSV M and S segments found in tobacco (*Nicotiana tabacum*) in SA

Segment	Segment Length(nt)	ORF position(nt)	ORF length(nt)	Polarity	Protein coded	Protein MW (kDa)	No. of Amino acids
L (GRSV)	8625	1-8625	8625	-	RdRp	330.39	2875
M (TCSV)	4814	101-1011	910	+	NSm	33.93	304
		4764-1360	3404	-	Gn-Gc	127.69	1135
S (TCSV)	3311	88-1497	1409	+	NSs	52.54	470
		3160-2384	776	-	N	28.56	259

3.3.5. Phylogenetic analysis

3.3.5.1. Tomato spotted wilt virus

The phylogenetic analysis was conducted using closely related TSWV nucleotide sequences for S, M, and L segments from NCBI, using tomato leaf curl New Delhi virus (ToLCNDV) as the outgroup to root the trees (Table 3.6). Our results show that the L segment of TSWV_Nto isolate (from this study) clustered with the LK-1 isolate from SA with the bootstrap value of 95% (Fig. 3.1), and it also had a nucleotide identity of 96.24% with the same isolate (Table 3.3). The M segment of TSWV did not cluster with any of the isolates used in the study (Fig. 3.2). The S segment TSWV clustered with isolate Mum-A5 from Zimbabwe and formed a distinct clade with isolates LK-1 from South Africa and QLD1 & 2 from Australia (Fig. 3.3).

Table 3.6 Complete genome and partial gene sequences of TSWV obtained from the NCBI nucleotide database used for phylogenetic analysis in conjunction with the complete sequences of the L, M, and S of the TSWV isolate from this study and ToCLNDV as an outgroup

Virus	Segment	ISOLATE	ORIGIN	ACCESSION NO	REFERENCE
ToLCNDV		-	Indonesia	LC431620	Unpublished
TSWV	S	TSWV_Nto	SA	OL471721	This study
TSWV		LK-1	SA	KY250488	Unpublished
TSWV		Mum-A5	Zimbabwe	MG602673	Karavina *et al.,* 2020
TSWV		DSMZ PV-1082	Bulgaria	MZ202330	Unpublished
TSWV		DSMZ PV-0393	Bulgaria	MW854273	Unpublished
TSWV		-	Bulgaria	D13926	De Haan *et al.,*1990
TSWV		DH37	Bulgaria	AJ418777	Heinze *et al.,* 2003
TSWV		BS97	Bulgaria	AJ418779	Heinze *et al.,* 2003
TSWV		HUP4-2012-WT	Hungary	KJ649611	Almasi *et al.,* 2015
TSWV		HUP5-2009-WT	Hungary	KJ649612	Almasi *et al*., 2015
TSWV		Ka-To	India	MK97765	Unpublished
TSWV		GD98	Bulgaria	AJ418780	Heinze *et al.,* 2003
TSWV		TSWV-QLD2	Australia	MG025804	Unpublished
TSWV		TSWV-QDL1	Australia	KT17693	Unpublished
TSWV		CPNH9	Brazil	NC002051	de Haan *et al.,* 1990
TSWV		p170	Italy	DQ431237	Unpublished
TSWV		NC-3	USA	AY744478	Tsompana *et al.,* 2005
TSWV		NC-1	USA	AY744476	Tsompana *et al.,* 2005
TSWV		p202	Italy	DQ398945	Unpublished
TSWV		p202/3WT	Italy	HQ830187	unpublished
TSWV		p202/3RB	Italy	HQ830186	Margaria *et al*., 2014
TSWV		TSW-D	USA	AF020660	Qiu *et al.,* 1998
TSWV		TSWV-12	South Korea	KC261967	Lian *et al.,* 2013
TSWV		TSWV-17	South Korea	KC261973	Lian *et al*., 2013
TSWV		LE98/257	Germany	AJ418781	Heinze *et al.,*2003
TSWV		Rujol1tl3	Spain	KP008131	Debreczeni *et al.*,2015
TSWV		TSWVantRB	Turkey	MH367502	Fidan & Sari 2019
TSWV		PepCal12	Italy	MG989676	Unpublished
TSWV		PepCal21	Italy	MG989676	Unbublished
TSWV	M	TSWV_Nto	SA	OL471722	This study

Virus	Segment	ISOLATE	ORIGIN	ACCESSION NO	REFERENCE
TSWV		NC-3	USA	AY744486	Tsompana *et al.,* 2005
TSWV		Cr1NL2	Spain	HM015511	Lopez *et al.,* 2011
TSWV		TSWV-17	South Korea	KC261972	Lian *et al.,* 2013
TSWV		P202/3RB	Italy	HQ830185	Margaria *et al.,*2014
TSWV		TSWV-12	South Korea	KC261966	Lian *et al.,*2013
TSWV		Da1NL2	Spain	HM015512	Lopez *et al.,* 2011
TSWV		Mum-A5	Zimbabwe	MG602672	Karavina *et al.,* 2020
TSWV		Pulou1TL3	Spain	HM015520	Debreczeni *et al.*,2015
TSWV		Oller1TL3	Spain	HM015519	Lopez *et al.,* 2011
TSWV		TSWVAntRB	Italy	MH367503	Margaria *et al.,*2014
TSWV		D	Netherlands	AF208497	Hoffman *et al.,*2001
TSWV		Ag1TL3	Spain	HM015515	Lopez *et al.,* 2011
TSWV		Ber1TL3	Spain	HM015517	Lopez *et al.,* 2011
TSWV		Sala1TL3	Spain	HM015521	Lopez *et al.,* 2011
TSWV		LYE51	France	MK792775	Unpublished
TSWV		WA7	Australia	KM365065	Wylie *et al.*, 2015
TSWV		D-191	Spain	HM015516	Lopez *et al.,* 2011
TSWV		SPAIN-2	Spain	AY744493	Tsompana *et al.,* 2005
TSWV		14-100	USA	KU179576	Unpublished
TSWV		LK-1	SA	KY250489	Unpublished
TSWV		TSWV-QLD1	Australia	KT717692	Unpublished
TSWV		TSWV-QLD2	Australia	MG025803	Unpublished
TSWV		Ka-To	India	MK977649	Unpublished
TSWV		DMSZ PV 0389	Bulgaria	MT723986	Unpublished
TSWV		DMSZ PV 0393	Bulgaria	MW854272	Unpublished
TSWV	L	TSWV_Nto	SA	OL471723	This sudy
TSWV		Mum –A5	Zimbabwe	MG602671	Karavina *et al.,* 2020
TSWV		WA7	Australia	KM365064	Wylie *et al.,* 2015
TSWV		p202/3WT	Italy	KJ575619	Unpublished
TSWV		DSMZ PV-0204	Bulgaria	MT682292	Unpublished
TSWV		LK-1	SA	KY250488	Unpublished
TSWV		TSWV-QLD1	Australia	KT717691	Unpublished
TSWV		Wa1	USA	MH745361	Unpublished
TSWV		BasC	USA	MH745369	Unpublished
TSWV		TSWV-4	South Korea	MH745370	Unpublished

Virus	Segment	ISOLATE	ORIGIN	ACCESSION NO	REFERENCE
TSWV		Tomato NJ-JN	South Korea	HM581934	Lee *et al.,* 2011
TSWV		TSWV-6	South Korea	KC261953	Lian *et al.,* 2013
TSWV		TSWV-7	South Korea	KC261957	Lian *et al.,* 2013
TSWV		TSWV-8	South Korea	KC261960	Lian *et al.,* 2013
TSWV		TSWV-10	South Korea	KC261962	Lian *et al.,* 2013
TSWV		TSWV-5	South Korea	KC261952	Lian *et al.,* 2013
TSWV		TSWV-16	South Korea	KC261969	Lian *et al.,* 2013
TSWV		TSWV-12	South Korea	KC261967	Lian *et al.,* 2013
TSWV		LYE51	France	MK792774	Unpublished
TSWV		TSWV 17	South Korea	KC261965	Unpublished
TSWV		pepCal24	Italy	MH763623	Unpublished
TSWV		pepCal22	Italy	MK348941	Unpublished
TSWV		Pepcal 10	Italy	MH763621	Unpublished
TSWV		LL—N05	Spain	KP008128	Debreczeni *et al.*, 2015
TSWV		Pujol1TL3	Spain	KP008130	Debreczeni *et al.*, 2015
TSWV		LN-HJL	China	MT241883	Unpublished
TSWV		NS BB20	South Korea	MT643236	Unpublished

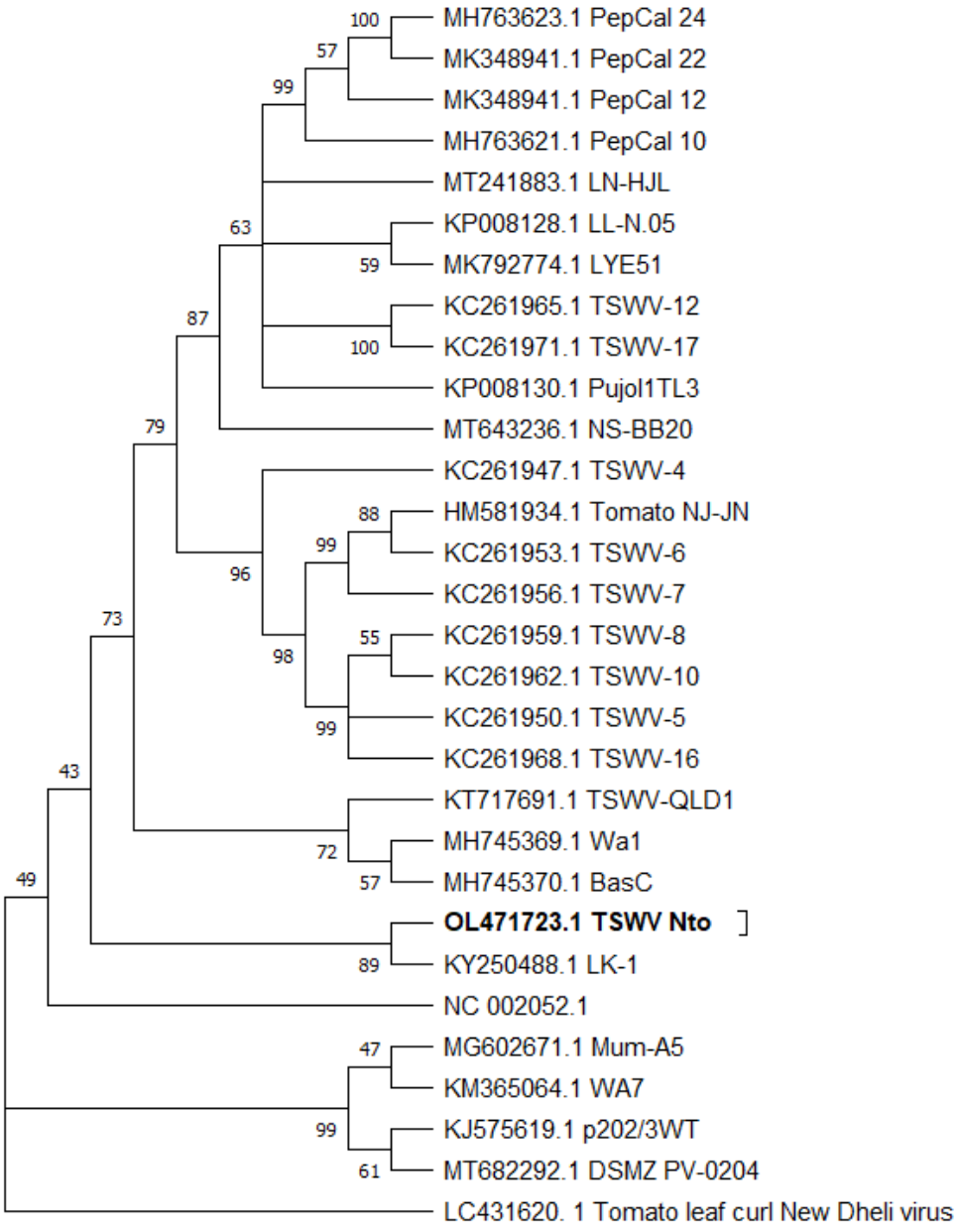


Figure 3.1: Phylogenetic analysis of the TSWV L segment isolate; TSWV_Nto (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

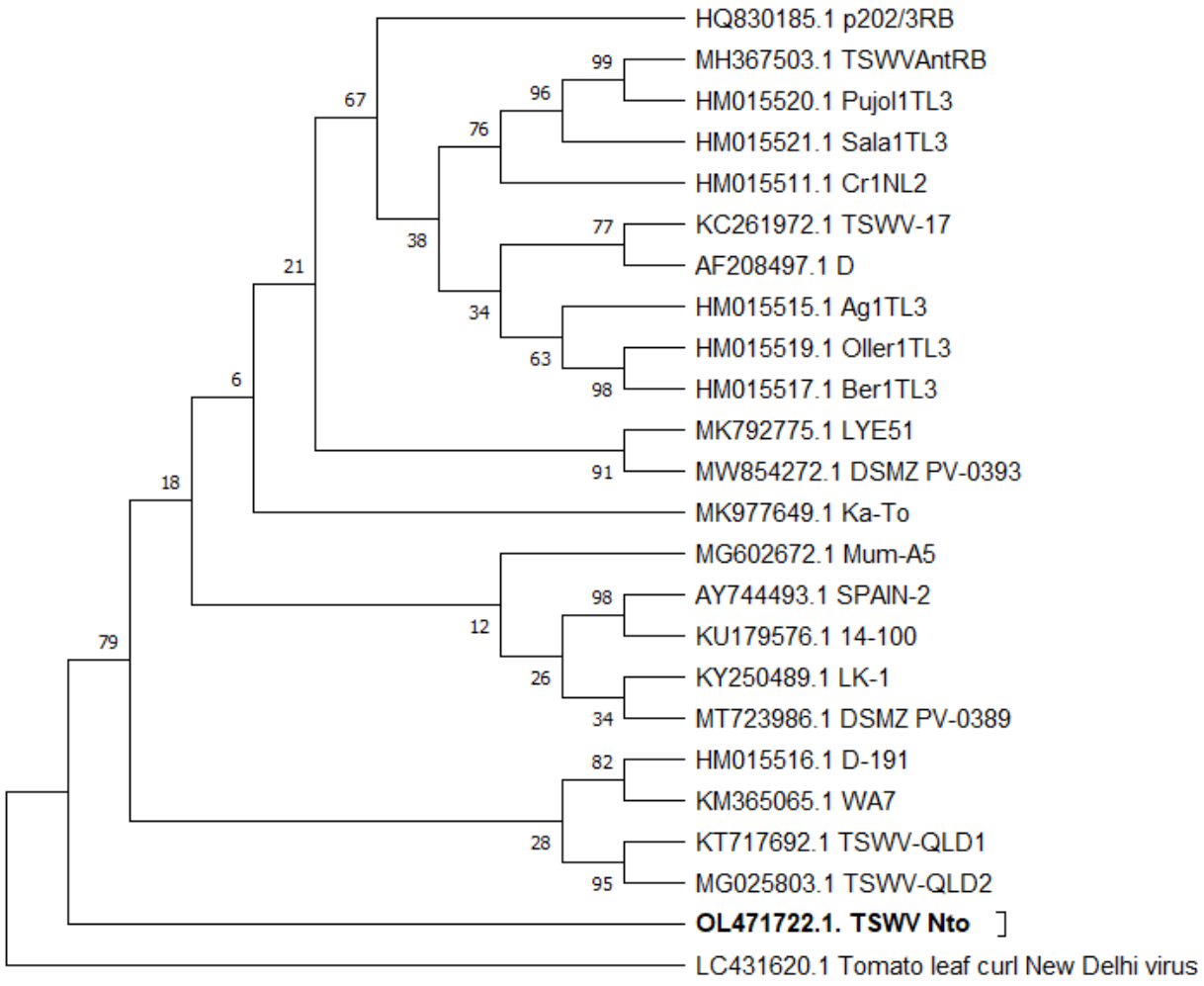


Figure 3.2: Phylogenetic analysis of the TSWV M segment isolate; TSWV_Nto (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

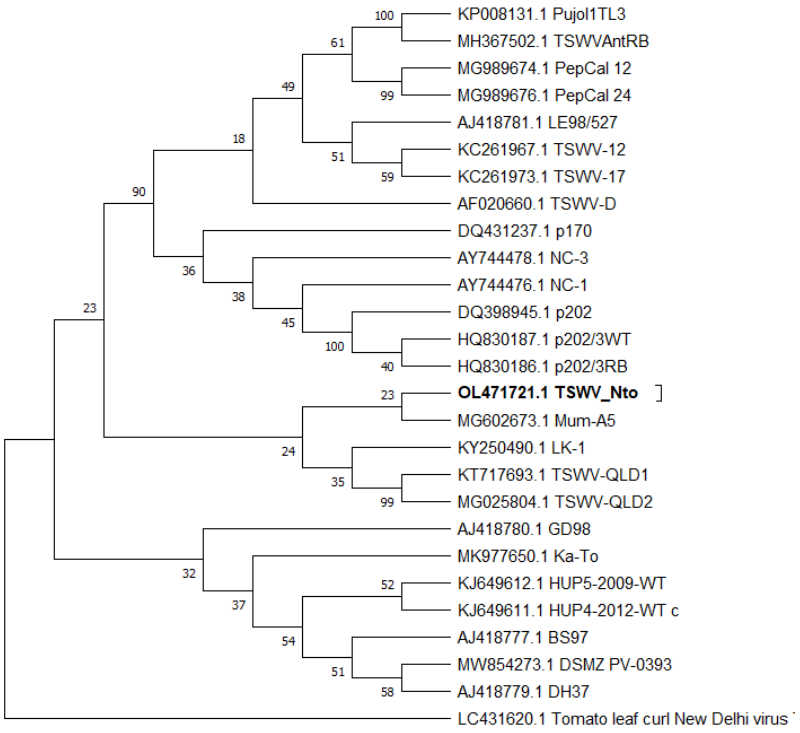


Figure 3.3: Phylogenetic analysis of the TSWV S segment isolate; TSWV_Nto (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

3.3.5.2. Groundnut ringspot virus and tomato chlorotic spot virus

The nucleotide sequences for each genome segment of GRSV and TCSV were subjected to the blastn tool to determine their closely related sequences. The phylogenetic analysis was conducted using closely related TCSV and GRSV nucleotide sequences for S, M, and L segments from the NCBI database. ToLCNDV was used as an outgroup to root the trees (Table 3.7). The closely related sequences were then used for the phylogenetic analysis. GRSV L segment from this study, isolate BonNDSA_GRSV (accession no. OK493952), did not cluster with any of the isolates (Fig. 3.4). TCSV M segment isolate BonNDSA_TCSV (accession no. OK493953) from this study clustered with the DR isolate from the Dominican Republic with a bootstrap value of 86%. (Fig. 3.5). TCSV S segment isolate BonNDSA_TCSV from this study (accession no. OK493954) formed a distinct clade with the two South African isolates SA-05, and DSMZ PV-0205 isolate (Fig. 3.6).

Table 3.7: Complete genome and partial gene sequences of TCSVand GRSV obtained from the NCBI nucleotide database used for phylogenetic analysis in conjunction with the complete sequences of the L (GRSV), M and S of the TCSV isolate from this study and ToCLNDV as an outgroup

VIRUS	SEGMENT	ISOLATE	ORIGIN	ACCESSION NO.	REFERENCE
TCSV	M	BonNDSA_TCSV	SA	OK493953	This study
TCSV		BR-03	Brazil	MH742960	Silva *et al.*,2019
TCSV		-	Brazil	AY574054	Resende et al., 2004
GRSV		LEM	Brazil	MH686227	De Marchi *et al.,* 2019
Reassortant		LMTSG	USA	HQ644141	Webster *et al.,* 2011
TCSV		DR	Dominican Republic	KX463273/NC035484	unpublished
GRSV		Haiti	Haiti	MH427862	Adegbola *et al.,* 2019
GRSV		BR	Brazil	KY350137	Unpublished
GRSV		GRSV-BR-PI	Brazil	MK503850	Unpublished
GRSV		SA-05	SA	AY574055	Resende *et al.*,2004
GRSV		GRSV-AR	Argentina	NC 043502	De Breuil *et al.*,2016
GRSV		-	Brazil	MN364669	Unpublished
GRSV		DSMZ PV 0205	SA	MT723992	Unpublished
TCSV	S	BonNDSA_TCSV	SA	OK493954	This study
TCSV+GRSV		LMTSG	USA	HQ644140	Webster *et al.,* 2011
GRSV		LEM	Brazil	MH686229	De Marchi *et al.,* 2019
GRSV		GRAR-S	Brazil	KY400110	Unpublished
Unknown		Unknown tospovirus	USA	L12048	Pang *et al.,* 1993
GRSV		SA-05	SA	MN364668	Silva *et al.*,2019
GRSV		DSMZ PV-0205	SA	MH742958	unpublished
TCSV		BR-03	Brazil	MT723993	Silva *et al.,* 2019
GRSV		Haiti	Haiti	MH427863	Adegbola *et al.,* 2019
GRSV		GRSV-AR	Argentina	NC 035484	De Breuil *et al.*,2016
TCSV		DR	Dominican Republic	KX463274	Unpublished
GRSV		UFT/2014	Brazil	MG797643	Unpublished
GRSV		BonNDSA_GRSV	SA	OK493952	This study
GRSV+TCSV	L	LSMTG	Brazil	HQ644142	Webster *et al.,* 2011
GRSV		Dsmz pv 0205	SA	MT723991	Unpublished
GRSV		LEM	Brazil	MH686228	De Marchi *et al.,* 2019
GRSV		GRSV-BR-PI	Brazil	MK503849	Unpublished
TCSV		DR	Dominican Republic	KX463272	Unpublished
GRSV		Haiti	Haiti	MH427861	Adegbola *et al.*,2019
TCSV		AB8900-2-3	Brazil	HQ700667	de Oliveira
TCSV		BR-O3	Brazil	MH742959	Silva *et al.,* 2019
GRSV		BR	Brazil	KY350136	Unpublished
GRSV		SA-05	SA	MH742956	Silva *et al.,* 2019
GRSV		-	Brazil	MN364670	unpublished
ToLNDV		ToLNDV	Indonesia	LC431620	Unpublished

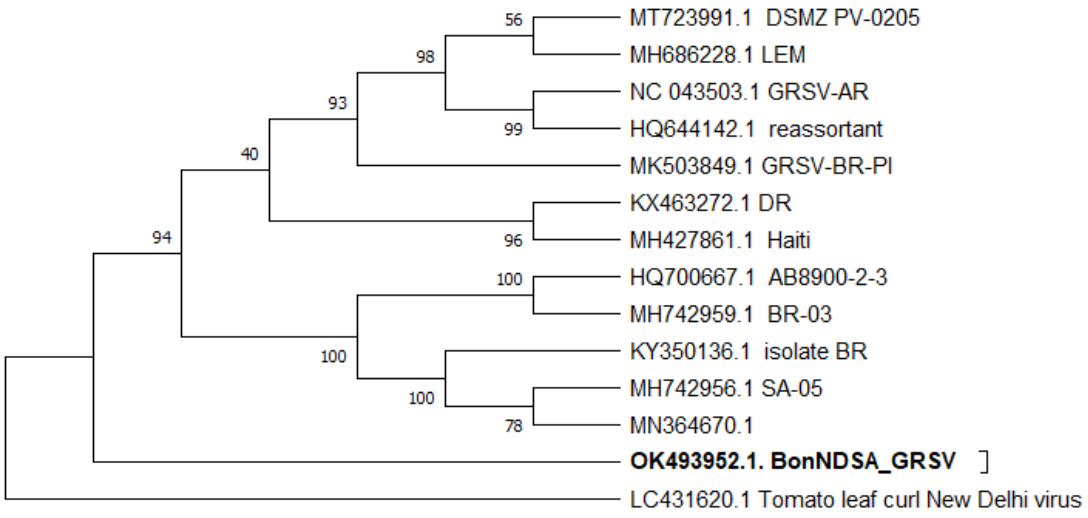


Figure 3.4: Phylogenetic analysis of the GRSV Lsegment isolate; BonDSA_GRSV (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

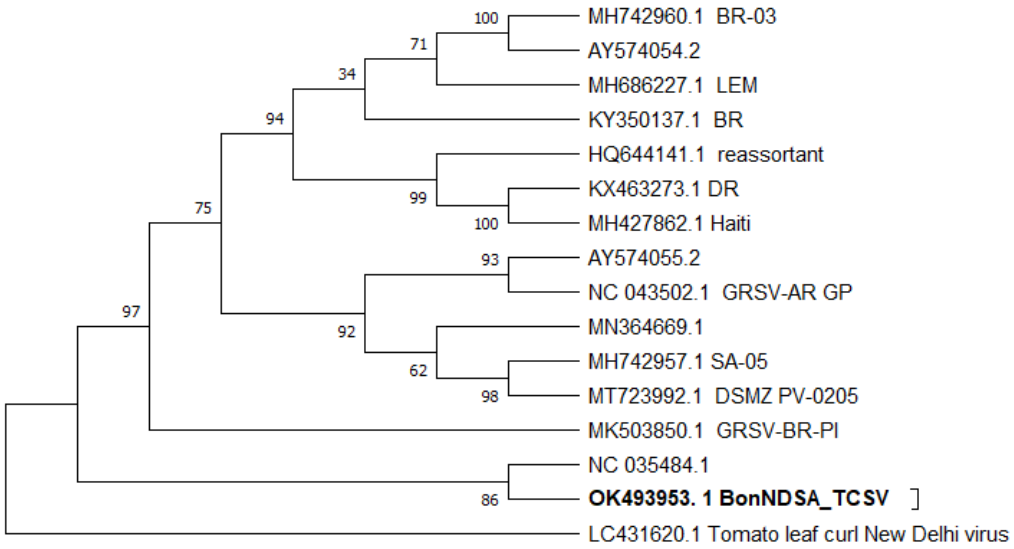


Figure 3.5: Phylogenetic analysis of the TCSV M segment isolate; BonDSA_TCSV (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

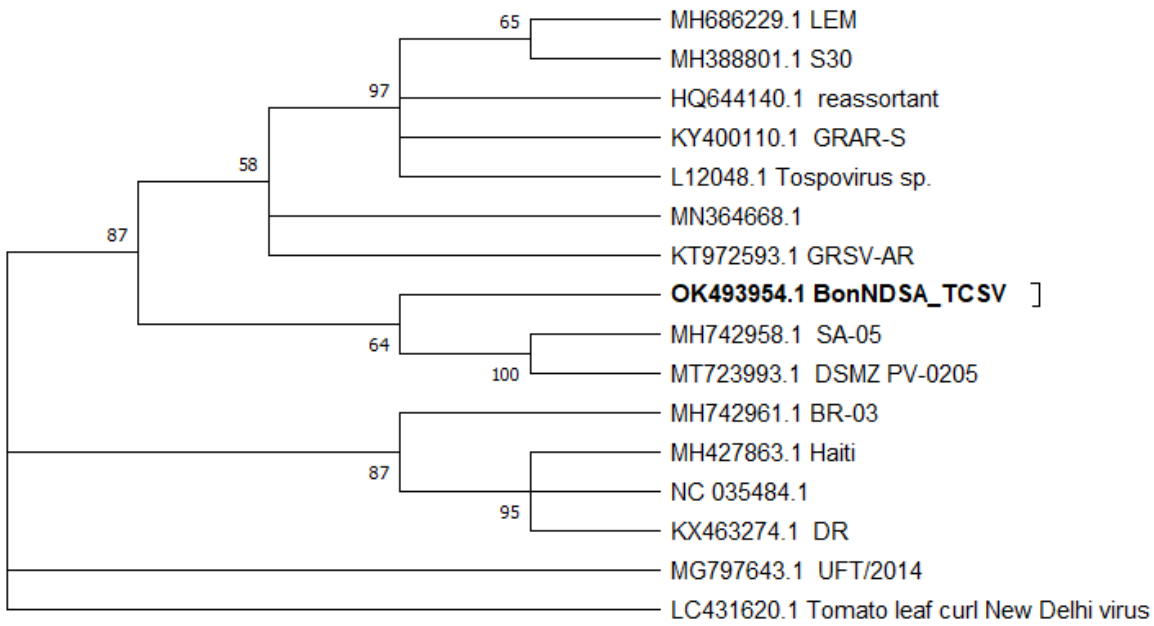


Figure 3.6c: Phylogenetic analysis of the TCSV S segment isolate; BonDSA_TCSV (Printed in bold on the phylogenetic tree) infecting *Nicotiana* sp. in South Africa using ToLCNDV as an outgroup. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The bootstrap consensus tree was inferred from 1000 replicates.

3.3.6. Recombination analysis

Sequences of the isolates under this study together with those in Table 3.8 were assessed for evidence of recombination using RDP4 software. Recombination events were detected only in three virus genomes: PVY, BMV, and ToLCUV. Each event was identified by at least four of the statistical programs in the RDP4 software. The PVY isolate Nto_PVY was involved in four recombination events as a minor parent the three Zimbabwean isolates; to d7-Bindura with 92.1% similarity (the event occurred at 420nt to 1636 nt), q19-Harare with 93.3% similarity (the event occurred at 452nt to 1494nt), and pep-Zim2 with 91.8% similarity (occurring at 471nt to 1636nt). Nto_PVY was a recombinant isolate of PVYNTN17 as a major parent, and SS607 31 as a minor parent with 88.1% similarity; the event occurred at 2254nt to 3193nt (Fig. 3.7).BMV isolate BMVt_SA was involved in one recombination event that occurred at 578nt to 861nt withs; BtMV-Wa from the USA as a major parent with 94.7% similarity and Inner Mongol from China as a minor parent with 86.3% similarity (Fig. 3.8). The ToLCUV isolate (from this study) was involved in two recombination events; ToLCUV_SA was a recombinant isolate of FLATZ15A9Sph C6, the

isolate from Tanzania as a major parent with 81.6% similarity and an unknown minor parent, the event occurred at 2626nt to 2672 (Fig 3.9). ToLCUV_SA was also a minor parent to isolate AFTZ15A6-6 with 89.1% similarity, the recombination occurred at 2455nt to 2524nt (Fig 3.9).

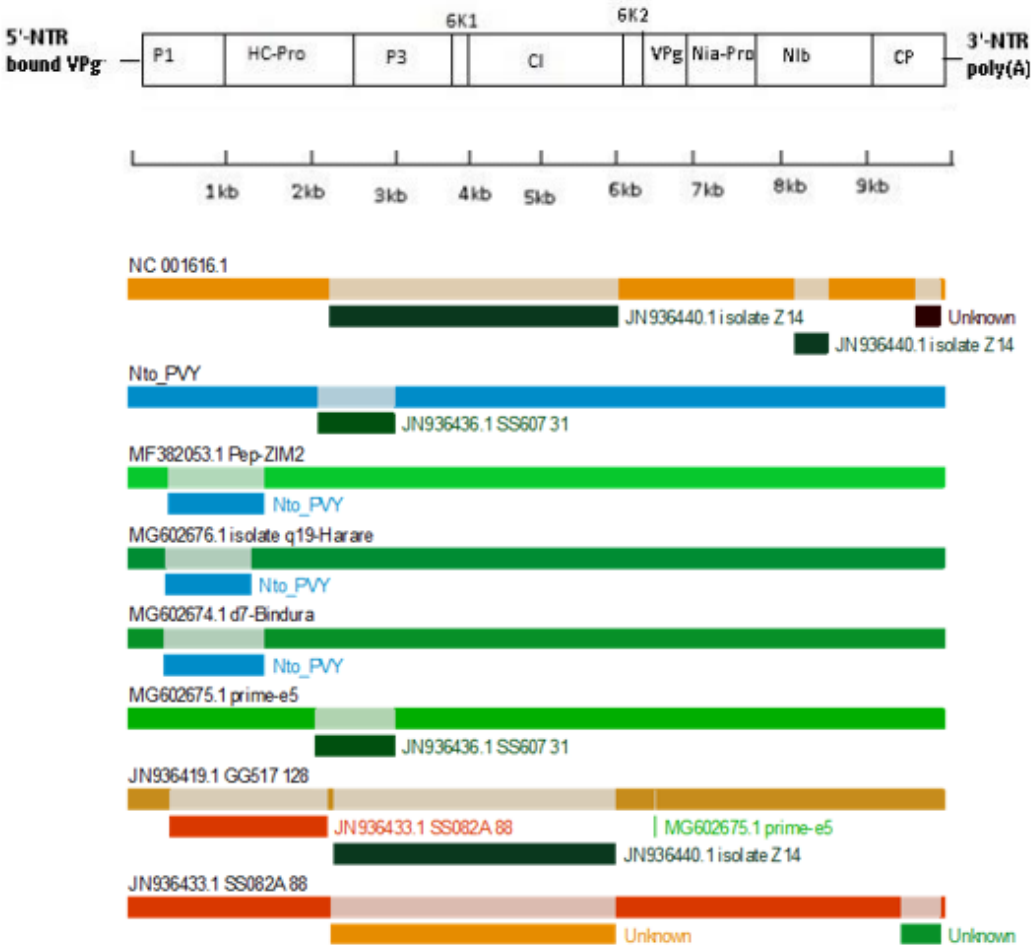


Figure 3.7: Recombination events occurring at one breakpoint of the sequence alignment for the SA PVY isolate (Nto_PVY) identified using RDP4 analysis. The top boxes indicate the various genes of the PVY polyprotein and indicate where exactly on the genome the different recombination events are occurring, respectively.

Table 3.8: Complete genome and partial genome sequences obtained from the NCBI nucleotide database used for recombination analysis in conjunction with the complete sequences of PVY, BMV and ToCLUBV isolates from this study

Virus	Isolate	Origin	Accession no	Reference
PVY	Z14	SA	JN936440	Unpublished
	SS607 31	SA	JN936436	Unpublished
	Pep-Zim2	Zimbabwe	MF382053	Unpublished
	q19-Harare	Zimbabwe	MG602676	Unpublished
	d7-Bindura	Zimbabwe	MG603674	Unpublished
	Prime-e5	Zimbabwe	MG603675	Unpublished
	GG517 12 8	SA	JN936419	Unpublished
	SS082A88	SA	JN936433	Unpublished
BMV	Inner Mongolia	China	DQ674263	Xiang *et al.,* 2007
	BtMV-Wa	USA	NC005304	Nemchinov *et al.,* 2004
	DSMZ PV1173	France	MW961164	Unpublished
	DSMZ PV 1228	Germany	MT815987	Unpublished
	DSMZ PV 1218	UK	MT815986	Unpublished
	DSMZ PV 1169	Sweden	MT737808	Unpublished
	DSMZ PV 0065	-	MT663309	Unpublished
ToCLUV	FLATZ15A9Bam_A6	Tanzania	MN381114	Unpublished
	-	Kenya	LC437667	Kimathi *et al.,* 2020
	FTZ15A6-6	Tanzania	MN381112	Unpublished
	FLATZ15A9Sph_C6	Tanzania	MN381113	Unpublished
	Ingaga	Uganda	DQ127170	Shih *et al.,* 2006
	FLATZ016Bam-I12	Tanzania	MN381115	Unpublished

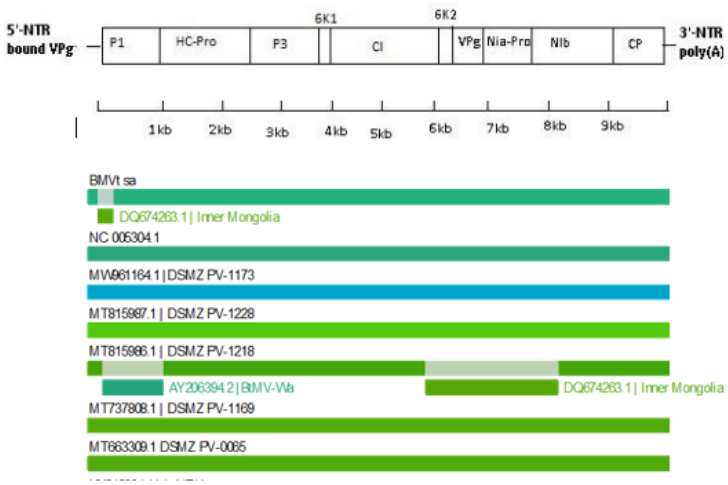


Figure 3.8.: Recombination events occurring at one breakpoint of the sequence alignment for the SA BMV isolate BMVt_sa identified using RDP4 analysis. The top boxes indicate the various genes of the BMV polyprotein and indicate where exactly on the genome the different recombination events are occurring, respectively.

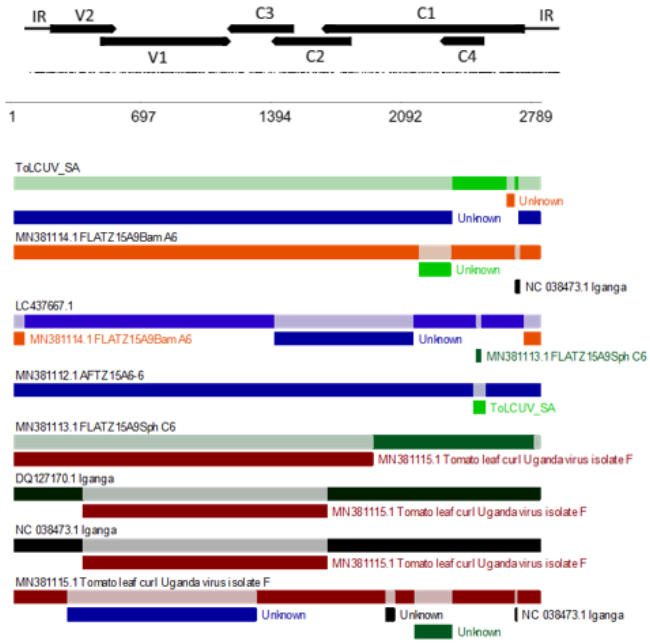


Figure 3.9: Recombination patterns occurring at two breakpoints of the sequence alignment for the SA ToLCUV_SA isolate identified using RDP4 analysis. The top lines indicate the various genes of the ToLCUV genome and indicate where exactly on the genome the different recombination events are occurring, respectively.

3.4. DISCUSSION

Identifying viruses is challenging based on the symptoms and using only specific methods such as RT-PCR, ELISA, and basic sequencing. Viruses elicit similar symptoms such as necrosis, chlorosis, mottling, deformation, and ringspot; NGS is one of the methods that can identify viruses without prior knowledge, i.e., symptoms and vector present. The results presented in this study showed that fifteen plant viruses (Table 3.2) are infecting tobacco in SA, including TSWV and GRSV that were identified and confirmed in Chapter 2 using ELISA, RT-PCR, and Sanger sequencing. NGS was able to recover both complete sequences and partial genome sequences.

When looking at the previous studies of viruses infecting tobacco from the past two decades, it appears that tobacco leaf curl disease has been a problem in SA. All these studies focused on the infections caused by viruses in the *Geminiviridae* family. The viruses identified in these studies were PVY, tobacco leaf curl virus (TLCV), tobacco apical stunt virus (TbASV), wound tumor virus (WTV), CMV, African cassava mosaic virus (ACMV), (Paximidis *et al.,* 1999; Rey *et al.,* 1999; Swanepoel and Net, 1995; and Rey 2020). No studies have used NGS to identify viruses in mixed infections in tobacco. Identification was based on the symptoms, using techniques such as ELISA, Bioassays, and RT-PCR. Among the viruses identified in this study, PVY, CMV, TMV, and TSWV have been reported on tobacco in SA in previous studies, and these viruses have been a concern in SA over the years (Moore, 1933; Paximidis *et al.,* 1999; Rey *et al.,* 1999; Swanepoel and Net, 1995; and Rey 2020). These viruses belong to the top 10 economically important plant viruses (Rybicki, 2015). Until recently, when orthotospovirus symptoms and thrips population became widely spread in the tobacco fields (discussed in Chapter 2).

The results presented in this study show that there are ten plant viruses and one satellite reported for the first time in SA. These viruses are; WAAV1, TCLV, PCVC, ToLCUV, BMV, BWYV, BCV-2, BCV-3, GRSV, TCSV, and cotton leaf curl Multan satellite. The viruses identified belonged to different genera; begomovirus, potyvirus, solendovirus, petuvirus, polerovirus, orthotospovirus, and deltapartitivirus. These results are an indication that different insect vector populations were present in the field, which may have caused mixed infection in the collected samples. Several factors contribute to the emergence of mixed plant virus infections, including genetic changes, recombination or reassortment in the virus genome, insect vector population changes, and migration due to crop trade and climate conditions changes (Dennehy, 2017).

NGS identified three Orthotospoviruses; TSWV, TCSV, and GRSV. TSWV and GRSV were already identified and confirmed using RT-PCR and ELISA, as described in Chapter 2. NGS recovered the whole genome of TSWV, one segment of GRSV, and two segments of TCSV. This study is one of a few to report the complete genome of orthotospovirus in SA. Only one study was done on TSWV by Kisten 2016, and one was done on GRSV by Silva *et al.* (2019).

Our study is the first to report GRSV and TCSV infecting tobacco in SA. TCSV M and S genome segments from this study share the closest nucleotide identity to two previously identified SA isolate, DSMZ PV 0205 and SA-05, respectively. From the study done by Silva et al. (2019), it appears that the SA-05 isolate is a parental genotype for both TCSV and GRSV. Genome rearrangement may have occurred between TCSV and GRSV, resulting in BonNDSA_TCSV being a reassortant isolate. The reassortment is common with orthotospoviruses in mixed infections (Martínez *et al.,* 2019); this is when part of the GRSV segment is putatively incorporated into the TCSV segment. According to recent studies, the evolution of orthotospoviruses is mainly caused by intra- and inter-specific viral genome rearrangements to adapt to different environmental conditions (Margaria *et al.,* 2007; Webster *et al.,* 2011, 2015; Tentchev *et al.,* 2011; Silva *et al.,* 2019).

The recombination analysis studies are important for the understanding of virus genomic evolution and hosts adaptations. The results presented in this study (Fig 3.7-3.9), together with other studies that have been done previously show that several of the isolates have undergone recombination as a result of mutation or genetic reassortment, as was previously observed in other isolates (Moodley *et al.,* 2014; Karavina *et al.,* 2021).

In conclusion, the complexity of virus populations identified in this study shows that it is important to test for viruses not only based on the visible symptoms or presence of vectors in the fields for proper disease monitoring strategies and efficient control strategies. NGS and bioinformatics pipelines are the essential tools in plant virology for time-efficient and virus identification without prior knowledge about the viruses infecting that particular crop.

3.5. REFERENCES

Adams, I. and Fox, A., 2016. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis. In *Current research topics in plant virology*: 323-335.

Adegbola, R.O., Kemerait, R.C., Adkins, S. and Naidu, R.A., 2019. Complete genome segment sequences of tomato chlorotic spot virus from peanut in Haiti. *Microbiology Resource Announcements*, *8*(18),: 306-319.

Akinyemi, I.A., Wang, F., Zhou, B., Qi, S. and Wu, Q., 2016. Ecogenomic survey of plant viruses infecting tobacco by next generation sequencing. *Virology Journal*, *13*(1): 1-12.

Almási, A., Csilléry, G., Csömör, Z., Nemes, K., Palkovics, L., Salánki, K. and Tóbiás, I., 2015. Phylogenetic analysis of Tomato spotted wilt virus (TSWV) NSs protein demonstrates the isolated emergence of resistance-breaking strains in pepper. *Virus Genes*, *50*(1): 71-78.

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

Barba, M., Czosnek, H. and Hadidi, A., 2014. Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses*, *6*(1): 106-136.

Bolger, A.M., Lohse, M. and Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15): 2114-2120.

Boni, M.F., Posada, D. and Feldman, M.W., 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*, *176*(2): 1035-1047.

Brown, J., Pirrung, M. and McCue, L.A., 2017. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, *33*(19): 3137-3139.

Buchfink, B., Xie, C. and Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1):59-60.

De Breuil, S., Cañizares, J., Blanca, J.M., Bejerman, N., Trucco, V., Giolitti, F., Ziarsolo, P. and Lenardon, S., 2016. Analysis of the coding-complete genomic sequence of groundnut ringspot virus suggests a common ancestor with tomato chlorotic spot virus. *Archives of Virology*, *161*(8):2311-2316.

De Haan, P., Wagemakers, L., Peters, D. and Goldbach, R., 1990. The S RNA segment of tomato spotted wilt virus has an ambisense character. *Journal of General Virology*, *71*(5): 1001-1007.

De Marchi, B.R., Bello, V.H., Watanabe, L.F.M., da Silva, F.B., Müller, C., Pavan, M.A. and Krause-Sakate, R., 2019. Characterization and complete genome sequence of groundnut ringspot orthotospovirus in soybean in Brazil. *Journal of Plant Pathology*, *101*(2): 401-401.

de Oliveira, A.S., Bertran, A.G.M., Inoue-Nagata, A.K., Nagata, T., Kitajima, E.W. and Resende, R.O., 2011. An RNA-dependent RNA polymerase gene of a distinct Brazilian tospovirus. *Virus Genes*, *43*(3): 385-389.

Debreczeni, D.E., López, C., Aramburu, J., Darós, J.A., Soler, S., Galipienso, L., Falk, B.W. and Rubio, L., 2015. Complete sequence of three different biotypes of tomato spotted wilt virus (wild type, tomato Sw-5 resistance-breaking and pepper Tsw resistance-breaking) from Spain. *Archives of Virology*, *160*(8): 2117-2123.

Dennehy, J.J., 2017. Evolutionary ecology of virus emergence. *Annals of the New York Academy of Sciences,* 1389(1): 124.

Felsenstein J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783-791.

Fidan, H. and Sarı, N., 2019. Molecular characterization of resistance-breaking Tomato spotted wilt virus (TSWV) isolate medium segment in tomato. *Applied Ecology and Environmental Research*, *17*(2): 5321-5339.

Flint, S.J., Racaniello, V.R., Rall, G.F., Hatziioannou, T. and Skalka, A.M., 2020. *Principles of virology, Volume 2: pathogenesis and control*. John Wiley & Sons.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D. and Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*: 571-607.

Gibbs, M.J., Armstrong, J.S. and Gibbs, A.J., 2000. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*, *16*(7): 573-582.

Heinze, C., Willingmann, P., Schwach, F. and Adam, G., 2003. An unusual large intergenic region in the S-RNA of a Bulgarian tomato spotted wilt virus isolate. *Archives of Virology*, *148*(1): 199-205.

Hoffmann, K., Qiu, W.P. and Moyer, J.W., 2001. Overcoming host-and pathogen-mediated resistance in tomato and tobacco maps to the M RNA of Tomato spotted wilt virus. *Molecular Plant-Microbe Interactions*, *14*(2): 242-249.

International Committee on Taxonomy of Viruses ICTV https://talk.ictvonline.org/ictv-reports/

Karavina, C., Ibaba, J.D. and Gubba, A., 2020. Full genome sequence of a chrysanthemum-infecting tomato spotted wilt virus isolate from Zimbabwe obtained by next-generation sequencing. *Acta Virologica*, *64*(1): 88-92.

Karavina, C., Ibaba, J.D. and Gubba, A., 2021. Potato virus Y isolates infecting bell pepper from parts of Southern Africa display distinct recombination patterns. *Physiological and Molecular Plant Pathology*, *114*: 101638.

Kimathi, R.H., Wilisiani, F., Mashiko, T., Neriya, Y., Miinda, A.E., Nishigawa, H. and Natsuaki, T., 2020. Molecular characterization of tomato leaf curl Uganda virus infecting tomato plants in Kenya. *Journal of General Plant Pathology*, *86*(6): 513-518.

Kisten, L., 2016. *Weeds: reservoirs of vector-borne viral pathogens in South Africa* (Masters dissertation) University of KwaZulu-Natal :55-83.

Kormelink, R., De Haan, P., Meurs, C., Peters, D. and Goldbach, R., 1992. The nucleotide sequence of the M RNA segment of tomato spotted wilt virus, a bunyavirus with two ambisense RNA segments. *Journal of General Virology*, *73*(11): 2795-2804.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* 35: 1547–1549

Leke, W.N., Mignouna, D.B., Brown, J.K. and Fondong, V.N., 2016. First report of Soybean chlorotic blotch virus and West African Asystasia virus 1 infecting cassava and a wild cassava relative in Cameroon and Togo. *New Disease Reports*, *33*: 24-24.

Lian, S., Lee, J.S., Cho, W.K., Yu, J., Kim, M.K., Choi, H.S. and Kim, K.H., 2013. Phylogenetic and recombination analysis of tomato spotted wilt virus. *PloS one*, *8*(5): 63380.

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. *Microbiology*, *81*(6): 1579-1585.

Lopez, C., Aramburu, J., Galipienso, L., Soler, S., Nuez, F. and Rubio, L., 2011. Evolutionary analysis of tomato Sw-5 resistance-breaking isolates of Tomato spotted wilt virus. *Journal of General Virology*, *92*(1):210-215.

Margaria, P., Bosco, L., Vallino, M., Ciuffo, M., Mautino, G.C., Tavella, L. and Turina, M., 2014. The NSs protein of tomato spotted wilt virus is required for persistent infection and transmission by Frankliniella occidentalis. *Journal of Virology*, *88*(10):5788-5802.

Margaria, P., Ciuffo, M., Pacifico, D. and Turina, M., 2007. Evidence that the nonstructural protein of tomato spotted wilt virus is the avirulence determinant in the interaction with resistant pepper carrying the Tsw gene. *Molecular Plant-Microbe Interactions*, *20*(5): 547-558.

Martin, D. and Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics*, *16*(6): 562-563.

Martin, D.P., Posada, D., Crandall, K.A. and Williamson, C., 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research & Human Retroviruses*, *21*(1): 98-102.

Martínez, R.T., de Almeida, M.M., Rodriguez, R., Cayetano, X., de Oliveira, A.S., Silva, J.M., Melo, F.L. and Resende, R.O., 2019. Analyses of orthotospovirus populations and dispersion under different environmental conditions in Brazil and in the Dominican Republic. *Tropical Plant Pathology*, *44*(6): 511-518.

Martínez, R.T., de Almeida, M.M.S., Rodriguez, R., de Oliveira, A.S., Melo, F.L. and Resende, R.O., 2018. Identification and genome analysis of tomato chlorotic spot virus and dsRNA viruses

from coinfected vegetables in the Dominican Republic by high-throughput sequencing. *Virology Journal*, *15*(1): 1-9.

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*: 90-96.

Moodley, V., Ibaba, J.D., Naidoo, R. and Gubba, A., 2014. Full-genome analyses of a Potato Virus Y (PVY) isolate infecting pepper (*Capsicum annuum* L.) in the Republic of South Africa. *Virus Genes*, *49*(3): 466-476.

Nemchinov, L.G., Hammond, J., Jordan, R. and Hammond, R.W., 2004. The complete nucleotide sequence, genome organization, and specific detection of Beet mosaic virus. *Archives of Virology*, *149*(6): 1201-1214.

Owen, J., Shintaku, M., Aeschleman, P., Tahar, S.B. and Palukaitis, P., 1990. Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 3. *Journal of General Virology*, *71*(10): 2243-2249.

Padidam, M., Sawyer, S. and Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology*, *265*(2): 218-225.

Pang, S.Z., Slightom, J.L. and Gonsalves, D., 1993. The biological properties of a distinct tospovirus and sequence analysis of its S RNA. *Phytopathology*, *83*(7): 728-733.

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.

Paximadis, M., Idris, A.M., Torres-Jerez, I., Villarreal, A., Rey, M.E.C. and Brown, J.K., 1999. Characterization of tobacco geminiviruses in the Old and New World. *Archives of Virology*, *144*(4): 703-717.

Posada, D. and Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences*, *98*(24): 13757-13762.

Qiu, W.P., Geske, S.M., Hickey, C.M. and Moyer, J.W., 1998. Tomato Spotted WiltTospovirusGenome Reassortment and Genome Segment-Specific Adaptation. *Virology*, *244*(1): 186-194.

Resende, R., Ávila, A. and Inoue-Nagata, A., 2004. Sequence analysis of the glycoproteins of Tomato chlorotic spot virus and Groundnut ringspot virus and comparison with other tospoviruses. *Virus Genes*, *29*(3): 321-328.

Rey, C., 2020. The hunt for plant viruses in South Africa: The early days. *South African Journal of Science*, *116*(11-12),: 1-8.

Rey, M.E.C., D'Andrea, E., Calvert-Evers, J., Paximadis, M. and Boccardo, G., 1999. Evidence for a phytoreovirus associated with tobacco exhibiting leaf curl symptoms in South Africa. *Phytopathology*, *89*(4): 303-307.

Rizzo, T.M. and Palukaitis, P., 1988. Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 2. *Journal of General Virology*, *69*(8): 1777-1787.

Rizzo, T.M. and Palukaitis, P., 1989. Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 1. *Journal of General Virology*, *70*(1): 1-11.

Rybicki, E.P., 2015. A Top Ten list for economically important plant viruses. *Archives of Virology*, *160*(1): 17-20.

Shih, S.L., Green, S.K., Tsai, W.S. and Ssekyewa, C., 2006. Molecular characterization of a begomovirus associated with tomato leaf curl disease in Uganda. *Plant Disease*, *90*(2): 246-246.

Silva, J.M.F., de Oliveira, A.S., de Almeida, M.M.S., Kormelink, R., Nagata, T. and Resende, R.O., 2019. Tomato chlorotic spot virus (TCSV) putatively incorporated a genomic segment of groundnut ringspot virus (GRSV) upon a reassortment event. *Viruses*, *11*(2): 187.

Smith, J.M., 1992. Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, *34*(2): 126-129.

Swanepoel, A.* & Nel, L., 1995. Epidemiological aspects of cucumber mosaic virus on flue-cured tobacco in South Africa. *African Plant Protection*, *1*(1): 13-18.

Tamura, K. and Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, *10*(3): 512-526

Tentchev, D., Verdin, E., Marchal, C., Jacquet, M., Aguilar, J.M. and Moury, B., 2011. Evolution and structure of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights into emergence processes. *Journal of general Virology*, *92*: 961-973.

Tsompana, M., Abad, J., Purugganan, M. and Moyer, J.W., 2005. The molecular population genetics of the Tomato spotted wilt virus (TSWV) genome. *Molecular Ecology*, *14*(1): 53-66.

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.

Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., Pauwels, R., Alcantara, L.C., Vanden Eynden, E., Vandamme, A.M. and Deforche, K., 2019. Genome Detective: an automated system for virus identification from high-throughput sequencing data. *Bioinformatics*, *35*(5): 871-873.

Webster, C.G., Frantz, G., Reitz, S.R., Funderburk, J.E., Mellinger, H.C., McAvoy, E., Turechek, W.W., Marshall, S.H., Tantiwanich, Y., McGrath, M.T. and Daughtrey, M.L., 2015. Emergence of Groundnut ringspot virus and Tomato chlorotic spot virus in vegetables in Florida and the southeastern United States. *Phytopathology,* 105(3): 388-398.

Webster, C.G., Reitz, S.R., Perry, K.L. and Adkins, S., 2011. A natural M RNA reassortant arising from two species of plant-and insect-infecting bunyaviruses and comparison of its sequence and biological properties to parental species. *Virology*, *413*(2): 216-225.

Whitfield, A.E., Ullman, D.E. and German, T.L., 2005. Tospovirus-thrips interactions. *Annual Review of Phytopathology* , *43*: 459-489.

Wylie, S.J., Zhang, C., Long, V., Roossinck, M.J., Koh, S.H., Jones, M.G., Iqbal, S. and Li, H., 2015. Differential responses to virus challenge of laboratory and wild accessions of Australian species of Nicotiana, and comparative analysis of RDR1 gene sequences. *PloS one*, *10*(3): 0121787.

Xiang, H., Han, Y.H., Han, C., Li, D. and Yu, J., 2007. Molecular characterization of two Chinese isolates of Beet mosaic virus. *Virus Genes*, *35*(3): 795-799.

Xie, W.S., Antoniw, J.F. and White, R.F., 1993. Nucleotide sequence of beet cryptic virus 3 dsRNA2 which encodes a putative RNA-dependent RNA polymerase. *Journal of General Virology*, *74*(7): 1467-1470.

Zaidi, S.S.E.A., Tashkandi, M., Mansoor, S. and Mahfouz, M.M., 2016. Engineering plant immunity: using CRISPR/Cas9 to generate virus resistance. *Frontiers In Plant Science*, *7*:1673.

Zhao, Y., Yang, X., Zhou, G. and Zhang, T., 2020. Engineering plant virus resistance: from RNA silencing to genome editing strategies. *Plant Biotechnology Journal*, 18(2): 328-336

Chapter 4: Dissertation overview

Tobacco is an economically important cash crop in South Africa (SA). Plant viruses are amongst the most important plant pathogens that cause a significant decrease in crop yield quality and economic losses worldwide including in SA (Zhao *et al.,* 2020). Previous studies on viruses infecting tobacco in SA during the past two decades have shown that tobacco leaf curl disease (TLCD) and a number of other viruses are a major limiting factor to economic production. However, these studies were not very successful in identifying some of the individual viruses that were associated with TCLD. Against this background, there is therefore need to look at modern molecular techniques to accurately identify some of the viruses infecting tobacco in SA.The aim of this study was to identify and characterize viruses infecting tobacco in major tobacco growing regions of South Africa. The first part of the study focused on identifying viruses based on the symptoms and associated insect vectors observed in the field using common detection methods; ELISA, RT-PCR, and Sanger sequencing. The second part focused on identifying viruses in mixed infections using next generation sequencing (NGS) and the characterization of the identified orthotospoviruses. NGS is a molecular technique which does not require prior knowledge about the virus and it is the useful tool for detecting viruses in mixed infections.

Major Findings

The techniques used in Chapter 2; ELISA, RT-PCR and Sanger sequencing were able identify and confirm the presence of two Orthotospoviruses; TSWV and GRSV in 19 of the 22 collected field samples from all the sampling sites. The NGS results in Chapter 3 confirmed the presence of three orthotospoviruses (TSWV, GRSV, and TCSV) in SA along with other 12 plant viruses from different genera; potyviruses (PVY and BMV), cucumovirus (CMV), begomovirus (WAAV1 and ToCLUV), solendovirus (TVCV), petuvirus (PVCV), tobamovirus (TMV), deltapartitivirus (BCV-2 and BCV-3) and polerovirus (BWYV). Nine of the identified viruses were reported for the first time in SA. Recombination events were observed in some of the isolates; BMVt_sa, Nto_PVY and ToCULV_sa. There were also about 40 other plant viruses detected by NGS with fewer reads and shorter partial nucleotide sequences (Data not presented). Most of these viruses

belonged to *Caulimoviridae* family. The results presented in this study highlight the importance of using NGS compared to traditional techniques used for virus identification.

Implications of findings

The results presented in this study confirmed the presence of orthotospoviruses using both specific detection methods and NGS. However, NGS was able to detect other viruses that were part of mixed infections. This highlights the importance of using NGS for identifying viruses not only based on the visible symptoms or presence of vectors in the fields for proper disease monitoring strategies and efficient control strategies. NGS can detect viruses rapidly even when in very low concentrations in an infected host. NGS can detect both DNA and RNA viruses, and it is also possible to assemble partial or complete virus sequences. NGS is one of the most accurate and time-efficient detection methods for virus identification (Akinyemi *et al.,* 2016).

There are complex factors associated with viral disease outbreaks, including rapid evolution, insect vector migration, and unpredictable increase of viral host range, make it very difficult to develop effective long-term disease management strategies (Zaidi *et al.,*2016). Several factors contribute to mixed plant virus infections and these include; genetic changes, recombination or assortment in the virus genome, insect vector population changes, and migration due to crop trade and changes of climate conditions (Dennehy, 2017).

Way forward

Efficient and accurate methods of virus identification are key towards the development of sustainable viral disease management strategies. Techniques such NGS should be used when virus identification studies are conducted on field samples, and each sample must be analysed separately in order to know the exact virus population in the sample provided there is enough funding. This approach would save time and money while generating a lot of useful information. The viruses that were identified in this study for the first time in SA by NGS should be subjected to further tests such as ELISA, biological assays, electron microscope and RT-PCR. Different inoculation techniques should be tested such as vector-mediated transmission or infectious clones. There should be more complete genome studies as this will assist with better understanding of the biology of these viruses at a molecular level. This knowledge can be used in the development of sustainable control strategies for virus diseases in general.

The samples used for this study were collected from the large-scale farming fields, which is a bit biased because 96 % of tobacco farmers in SA are small holder farmers (AgriSETA 2021). For future studies it would advisable that both commercial and small holder farmers be sampled when such studies are done. It is recommended that different tobacco varieties are assayed for sources of resistance or tolerance. Current circumstances indicate that it may require a study on pathogen-derived resistance using commercially cultivated varieties. There should be platforms where smallholder farmers are able to communicate with plant pathologists for the identification of unusual symptoms or insects vectors in their fields. This will go a long way towards helping this sector deal with diseases affecting their crops.

References

AgriSETA https://www.agriseta.co.za/wpcontent/uploads/2021/02/Agriseta_Tobacco_

Akinyemi, I.A., Wang, F., Zhou, B., Qi, S. and Wu, Q., 2016. Ecogenomic survey of plant viruses infecting tobacco by next generation sequencing. *Virology journal*, *13*(1): 1-12.

Dennehy, J.J., 2017. Evolutionary ecology of virus emergence. *Annals of the New York Academy of Sciences*, 1389(1): 124.

Diop, S.I., Geering, A.D., Alfama-Depauw, F., Loaec, M., Teycheney, P.Y. and Maumus, F., 2018. Tracheophyte genomes keep track of the deep evolution of the Caulimoviridae. *Scientific Reports*, *8*(1): 1-9.

Geering, A.D., Maumus, F., Copetti, D., Choisne, N., Zwickl, D.J., Zytnicki, M., McTaggart, A.R., Scalabrin, S., Vezzulli, S., Wing, R.A. and Quesneville, H., 2014. Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. *Nature Communications*, *5*(1): 1-11.

Zaidi, S.S.E.A., Tashkandi, M., Mansoor, S. and Mahfouz, M.M., 2016. Engineering plant immunity: using CRISPR/Cas9 to generate virus resistance. *Frontiers in Plant Science*, *7*:1673.

Zhao, Y., Yang, X., Zhou, G. and Zhang, T., 2020. Engineering plant virus resistance: from RNA silencing to genome editing strategies. *Plant Biotechnology Journal*, *18*(2): 328-336