Identification and Characterization of Viruses Infecting Soybean (*Glycine max*. L) in KwaZulu-Natal, South Africa

By

Nobuhle Ndlovu

(215030013)

BSc Agric (Hons)

Submitted in fulfillment of the requirements for the degree of

Master of Science

In the Discipline of Plant Pathology

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

December 2021

Dissertation Abstract

Soybean (*Glycine max* L.) is the world's most important seed legume, primarily used as an oil crop and protein source. Plant viruses are a major limiting factor to soybean production worldwide. Many destructive plant viruses have been discovered to infect soybean on a global scale. In South Africa, few viruses have been detected infecting soybean crops. The studies of viruses on soybean were undertaken several years ago and it is possible that the virus population structure may have evolved over time. The aim of the study was to identify and characterise viruses infecting some soybean cultivars grown in KwaZulu-Natal. Additionally, to determine the incidence of seed transmitted viruses.

The first part of the study was undertaken to detect and identify viruses presently infecting soybean grown in the province of KwaZulu-Natal, South Africa and determine the incidence of any seed transmitted viruses. Fifty-four soybean leaf samples exhibiting virus-like symptoms were collected from breeding lines growing in a Plant Pathology disease garden and greenhouses at the University of KwaZulu-Natal during the 2018 - 2019 and 2019 - 2020 growing seasons. Mechanical inoculation using inoculum prepared from the soybean field samples was done on *Nicotiana tabacum* L. to propagate the viruses in the collected samples. Symptom development was monitored on inoculated *N. tabacum* plants for 2-3 weeks after inoculation. The field samples were also subjected to Reverse Transcription Polymerase Chain Reaction (RT-PCR) and PCR to detect viruses known to infect soybean worldwide. Generic and specific primers were used to target specific coding regions of the viruses tested. Antibodies specific to cucumber mosaic virus (CMV) and tobacco mosaic virus (TMV) were used to test for virus presence in the field samples using double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA). The study also focused on determining the incidence of seed transmitted viruses by planting different soybean varieties/lines and testing for virus presence after the plants had germinated. Virus presence was based on symptoms exhibited by the germinated plants and by DAS-ELISA. Results of this study showed that the inoculated *N. tabacum* plants developed virus-like symptoms. Soybean mosaic virus (SMV), TMV, CMV, and hibiscus chlorotic ringspot virus (HCRSV) were identified in the field samples based on PCR results. Seed transmission assays did not demonstrate the presence of viruses based on symptomatology and DAS-ELISA tests.

The second part of the study was undertaken by using Next Generation Sequencing (NGS) to analyse the complete genome sequence of HCRSV infecting soybean in the province of KwaZulu-Natal, South frica Total RNA extracted from soybean samples exhibiting virus-like symptoms was combined into one sample and used as template for NGS analysis. The sequence data generated was analysed using Genome Detective Virus Tool version 1.133.

The HCRSV complete genome sequence obtained was compared with other HCRSV sequences from GenBank database using BLASTN. Pairwise and Multiple sequence alignments of the sequences were done using ClustalW tool available in MEGA X. Phylogenetic analysis was done using nine closely related HCRSV sequences including turnip crinkle virus (TCV) which was used as an outgroup. The open reading frames (ORFs) for the HCRSV genome were determined using ORF finder and protein sizes were measured using Protein Molecular Weight software. Recombination events were analysed using RDP4 software. NGS data analysis revealed that HCRSV, CMV and TMV were present in the infected soybean samples. Results from the phylogenetic analysis showed that the NdlovuNS_HCRSV-SA isolate from this study (Accession number: OK636421) was closely related to isolate XM from China with a bootstrap value of 99%. Genome organisation analysis of the NdlovuNS_HCRSV_SA isolate compared with other HCRSV isolates suggested high levels of similarity. The BLAST analysis correlated with the results from the genome organisation data, with the HCRSV isolates sharing 87.87% - 97.10% nucleotide identity. Recombination analyses showed a single event confirming that the NdlovuNS_HCRSV-SA isolate is a recombinant strain.

Accurate detection and identification of viruses plays an important role in virus disease management. Undetected viruses many occur and cause severe losses in soybean production. In this study, molecular detection techniques were used to accurately detectand identify the viruses infecting soybean field samples. It is important to emphasize that accurate and early detection of viruses is crucial for application of proper and effective control measures. The findings of this study will contribute to the body of knowledge on viruses infecting soybean in South Africa and will help in developing effective control measures.

Table of Contents

Declaration

Herein, I **Nobuhle Ndlovu** declare that:

a. This dissertation is my original work and has not been presented for a degree or examination in any other University.
b. This dissertation does not contain other persons' data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.
c. This dissertation does not have another person's writing, unless specifically acknowledged
d. This dissertation does not have text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

14/12/2021

Signature (Candidate) **Date**

14/12/2021

Signature (Supervisor) **Date**

Acknowledgements

I am very thankful to the Lord God almighty for His unmerited favour throughout my life. It has been a quite a journey, but words cannot begin to explain how grateful I am firstly, to be alive and in good health to complete this study. Secondly, to have had such a great opportunity come my way is a huge blessing, I really thank the LORD for everything and everyone who assisted me.

Proverbs 15:22 [AMP], 'Without consultation and wise advice, plans are frustrated, But with many counselors they are established and succeed.' To the following people, I would like to say:

- Many thanks to my supervisor, Prof. Gubba for his applaudable supervision, and remarkable sense of care for students. I appreciate all the amazing support, patience, encouragement, editing this work and all his great efforts for the success of this study.
- Many thanks to the South African Cultivar and Technology Agency (SACTA) for all their amazing support and funding my study.
- Many thanks to the UKZN SRC 2020/2021 who have done an outstanding job in assisting students in various ways.
- Many thanks to Dr. Ibaba for all his assistance towards the success of this study.
- Many thanks to all the Plant pathology students, and a special thanks to Bongeka and Peace Mwelase for all the amazing support, assistance, sharing of knowledge and good laughs.
- Many thanks to the plant pathology technicians, Busisiwe, Khulekani, Matthew, Sue, and her staff in the greenhouses at CERU for all the assistance they have provided.
- Many thanks to Bab' Skhakhane (who has his office closer to the virology lab) for always being friendly and giving good laughs.
- Many thanks to the stuff at Inqaba Biotech and ARC biotechnology Platform for all their services.
- From the Microscopy and Microanalysis Unit (MMU) department, many thanks to Ms. Cynthia and Dr. Beukes for their support and assistance
- From the department of Plant Breeding, many thanks to Dr. Shanahan, Prof. Sibiya, Dr. Musvosvi, and Lihle Mathonsi for assisting me with soybean seeds and samples.

- Many thanks to the Plant Pathology staff, Dr. Mbili, Dr. Yobo, and Dr. Bancole, for all their assistance and support when needed.
- Many thanks to Dr V. Moodley (Joey) for always willing to share information.
- Thanks to my home away from home, the Student Christian Fellowship (SCF), and Maritzburg Christian Church (MCC) for all their support, love, and warmth which they have consistently shown towards me.
- To everyone who have assisted me in one way or another, you are all greatly appreciated.
- Last but not least, many thanks go to my loving, prayerful family and friends who have consistently supported me since day one.

Dedication

I dedicate it to my grandmother (Mrs. C.O Mahlobo), my superwoman.

Introduction to Dissertation

Soybean (*Glycine max* L.) is known as a "miracle crop" that is annually grown in many parts of the world (Singh et al., 2019). It is mainly used as a vegetable protein and oil worldwide (Hartman et al., 2011). Soybean is the fourth largest grain or oilseed crop in the world after corn, wheat, and rice (Grain SA, 2019). The United States of America, Brazil, and Argentina are the world's major soybean producers, which account for 81% of global soybean production (Soystats, 2021; USDA, 2018). In South Africa, soybeans are mostly grown in Free State, Mpumalanga, and KwaZulu-Natal provinces (Soystats, 2021; DAFF, 2017). Interest in studying soybeans has lately surged in South Africa because of the health benefits linked with the crop and its products. Soybean farmers are constantly faced with many challenges in producing this crop, and the challenges include, among others, pests and diseases caused by viruses, fungi, and bacteria (Lal, 2009; Strange and Scott, 2005). Viruses are ranked as the second most important plant pathogens following fungi (Vidaver and Lambrecht, 2004). Plant virologists work tirelessly in order to accurately detect and identify the viruses present in various crops. This is done with the aim of developing effective and sustainable control strategies for these viruses and the diseases they cause.

Problem statement

Plant viruses are among the main limiting factors to soybean production (Hill and Whitham, 2014). Particularly seed-transmitted viruses have great potential to reduce yields because they interfere with the plant growth from the beginning (Sastry, 2013; Vroon et al., 1988). Yield losses can be up to 100% if the virus vectors are present, they rapidly spread the viruses to other healthy or uninfected plants. Mixed infections of viruses are very common and pose an even greater threat to soybean production. Viruses like soybean mosaic virus and bean pod mosaic virus act synergistically and induce severe symptoms resulting in reduction in soybean yields (Calvert and Ghabrial, 1983). The crop losses adversely affect human food including vegetable oil, seed-milk, and by-products, animal feed including chicken and pork, and biofuel (Zoundji et al., 2015).

Significance of Research

Limited studies have been undertaken on the identification of viruses infecting soybeans in South Africa in the last 10 years. Viruses evolve, so there is every likelihood that the virus population structure has changed over the years. Several viruses may be occurring in soybean in South Africa and remain unknown. This study will contribute to the body of knowledge on the identification and characterisation of viruses infecting soybean in South Africa and serves as the basis to develop effective control strategies.

The research objectives of this study were as follows:

- To detect and identify viruses presently infecting soybean grown in the province of KwaZulu-Natal, South Africa, and determine the incidence of any seed transmitted viruses.
- Use Next Generation Sequencing (NGS) to analyse the complete genome sequence of HCRSV infecting soybean in the province of KwaZulu-Natal, SA.

Dissertation Structure

The dissertation consists of four discrete chapters; Chapter 1 provides the review of current literature about the study, by outlining background information on soybean including the overview of soybean industry, the current status of soybean viruses in SA, detection methods currently used for soybean virus identification and control strategies used to manage the virus diseases. Chapter 2 focuses on the detection and identification of viruses infecting soybean (*Glycine max*. L) in KwaZulu-Natal province, South Africa. Chapter 3 concentrates on the complete genome analysis of hibiscus chlorotic ringspot virus isolate infecting soybeans in KwaZulu-Natal, South Africa. Then lastly, Chapter 4 gives the general overview of the study, which includes major findings, implications, and the way forward.

References

Calvert, L. A., and Ghabrial, S. A. (1983). Enhancement by Soybean Mosaic Virus of Bean Pod Mottle Virus titre in doubly infected soybean. Phytopathology, 73, 992-997.

DAFF (2017). Department of Agriculture, Fisheries and Forestry. Report 2017.11

Grain SA (2019). GROWING SOYBEANS, A great source of food, protein and oil. https://www.grainsa.co.za/growing-soybeans-a-great-source-of-food,-protein-and-oil. Accessed: 07 September 2020.

Hartman, G. L., West, E. D., and Herman, T. K. (2011). Crops that feed the World 2. Soybean worldwide production, use, and constraints caused by pathogens and pests. Food Security, 3(1), 5–17.

Hill, J. H., and Whitham, S.A. (2014). Control of Virus Diseases in Soybeans. Advances in Virus Research, 90, 355-390.

Lal, R. (2009). Soil degradation as a reason for inadequate human nutrition. Food Security, 1, 45–57.

Sastry, K.S. (2013). Seed-borne plant virus disease. Springer, New York. p. 328. First Edition.

Singh, G., Pradesh, M., and Biodiversity, S. (2019). 'A Review on Various Aspects of Soybean and Soybean Mosaic Virus', 1, 34-44.

SoyStats (2021). A reference guide to important soybean facts and figures. American Soybean Association. http://soystats.com/. Accessed: 13 July 2021

Strange, R. N., and Scott, P. R. (2005). Plant disease: a threat to global food security. Annual Review of Phytopathology, 43, 83–116.

USDA (2018). United States Department of Agriculture Report. Report 2018.

Vroon C. W., Pietersen, G., and Van Tonder, H. J. (1988). Seed transmission of soybean mosaic virus in *Lupinus albus* L. Phytophylactica, 20, 169-175.

Vidaver, A.K., and Lambrecht, P.A. (20,04). Bacteria as plant pathogens. The Plant Health Instructor.

Zoundji, C.C., Houngnandan, P., Dedehouanou, H., and Toukourou, F. (2015). Determinants of Soybean *Glycine max* (L.) production system in Benin. Journal of Experimental Biology and Agricultural Sciences, 3(5), 430-439.

Chapter 1

Literature Review

1.1. Introduction

Soybean (*Glycine max* L.), also called soya bean, is one of the most important crops worldwide (Kofsky et al., 2018). It is a popular and the most notable annual legume of the pea family, *Fabaceae* (Surekha et al., 2018). Soybean has a wide range of uses, but it is mainly used as a vegetable oil and a protein source (Singh et al., 2008). About 58% of total oilseed production worldwide and 69% of the world's protein concentrate for livestock feeding comes from soybean (Devine and Card, 2013; Chawla et al., 2013; Liu et al., 2016). Soybean proteins are also used for human consumption in many forms which include among others the soymilk for baby/infant formulas, flours for baking, and a meat substitute for vegetarians (Liu, 2008). Additional benefits associated with soybeans include their ability to fix nitrogen in the soil, which contributes to agriculture – intercropping and crop rotation purposes (Szostak et al., 2020). It is also known for its high tolerance for different soil and climate conditions (Szostak et al., 2020).

The origin of soybean was lost in obscurity due to limited molecular-based studies and archaeological information (Kofsky et al., 2018). Botanists believe it was first domesticated from the wild species *Glycine soja* Siebold & Zucc. in central China in the early 7000 BCE, estimated to be more than 5000 years ago (Guo et al., 2010; Li et al., 2013; Kim et al., 2010; Sedivy et al., 2017; Kofsky et al., 2018). In 1804, soybeans were introduced into the United States and Brazil (Kim et al., 2011), now regarded as the largest producers worldwide, followed by Argentina, China, and India (SoyStats, 2017). In South Africa (SA), soybean was newly introduced in 1903 (Dlamini et al., 2014; du Toit, 1942). Today, SA is among the major producers in Africa, followed by Nigeria, Zambia, Malawi, Benin, and Zimbabwe (Makurira, 2010). According to Statistica, the recent soybean production in SA by provinces showed that Free State, Mpumalanga and KwaZulu Natal are the leading producers. The overall soybean production in SA is estimated at between 100 000 and 800 000 tons per annum (DAFF, 2017).

Plant viruses are one of the main contributors to severe economic losses in a wide range of crops including soybean. Global economic losses of soybean due to viral diseases are estimated at several billion US dollars per year (Hartman et al., 2011). In SA, studies have made progress in identifying several viruses infecting soybean and the yield losses due to these viruses were recorded at up to 20% (Pietersen and Garnett, 1990; Pieterson et al., 1998). In another study, Lamprecht et al. (2010) reported a new Cytorhabdovirus, soybean blotchy mosaic virus (SbBMV) infecting soybean in SA. Given that these studies were

conducted many years back, it is possible that new viruses have emerged and remain unknown. Accurate and reliable detection methods are required to identify these viruses. This review gives background information on soybeans with a brief overview of the soybean industry, the current status of soybean viruses in SA, detection methods currently used for soybean virus identification, and the control strategies employed to manage the viral diseases.

1.2. Overview of Soybean Industry

1.2.1. Global Soybean Industry

The world's soybean production has largely increased over the last 20 years (Fig. 1.1). Soybean is considered a high-yielding crop which help in feeding the growing world population (Engelbrecht et al., 2020). Total soybean production was estimated at 358,8 million tons for the 2018/19 growing season, reported by the World Agricultural Supply and Demand Estimate (WASDE), 2021. In the 2019/2020 growing season, there was a 5,5% decline in the world's soybean production, estimated at 339,0 million tons. The decline was due to the impacts of the devastating COVID-19 pandemic which have indirectly affected the output on farms worldwide. Despite the negative impacts, there was an increase of 7,3% soybean production estimated at 361,0 million tons worldwide in 2020/2021 (Voora et al., 2020). The leading soybean producers in 2021, are Brazil with 38% contribution, followed by the United States with 29%, and Argentina contributing 14% (Fig. 1.2). As a collective, these countries account for 81% of the global soybean production (Soystats, 2021). According to the recent increase in the world soybean production, more production is expected in the upcoming years which will be matched with the rising demand for soybean products.

The largest soybean producers in the world remain globally competitive by implementing gradual processes of technological innovation (Figueiredo, 2016; FAO, 2017). The efforts made by crop scientists, plant pathologists, soybean growers, and many other role players lead to an increase in the world production of soybean (FAO, 2017). The package of technology includes the combination of direct seeding, plant nutrition, chemical control of weeds and diseases, efficient mechanical harvesting, the use of transgenic seeds, gradually improvingsoils, and use water more efficiently (Shea et al., 2020; OECD, 2001). In addition, soybean is easily rotated with other crops and cultivated in mixed and intercropping systems. One of the limiting factors for leading soybean producing countries is the availability of farmland due to more urbanization (OECD, 2001).

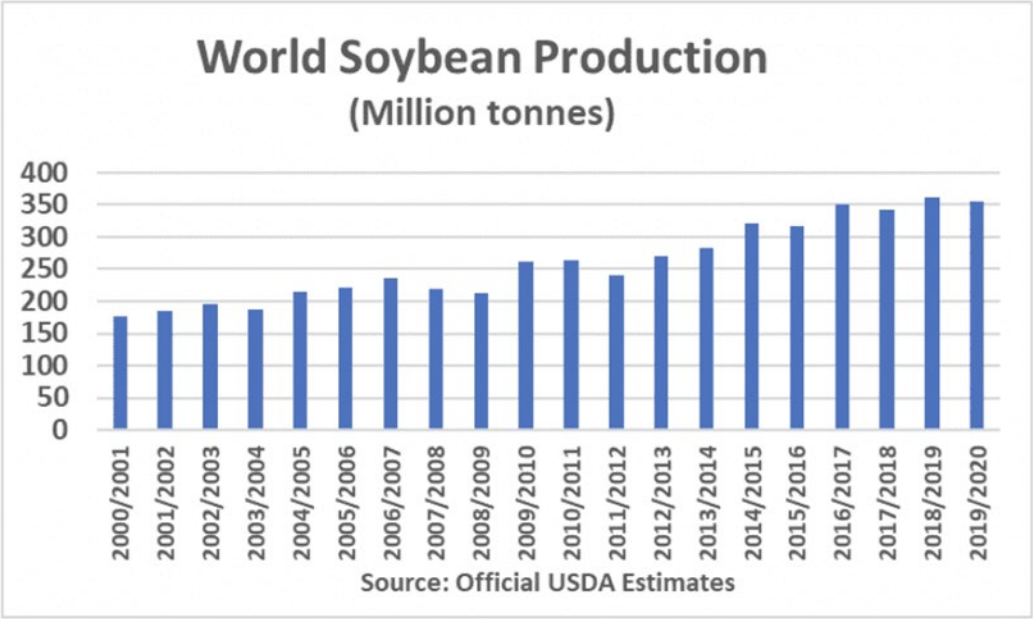

Figure 1.1: The Estimated World Soybean Production from 2000-2020 (Source: Official USDA Estimates, 2021)

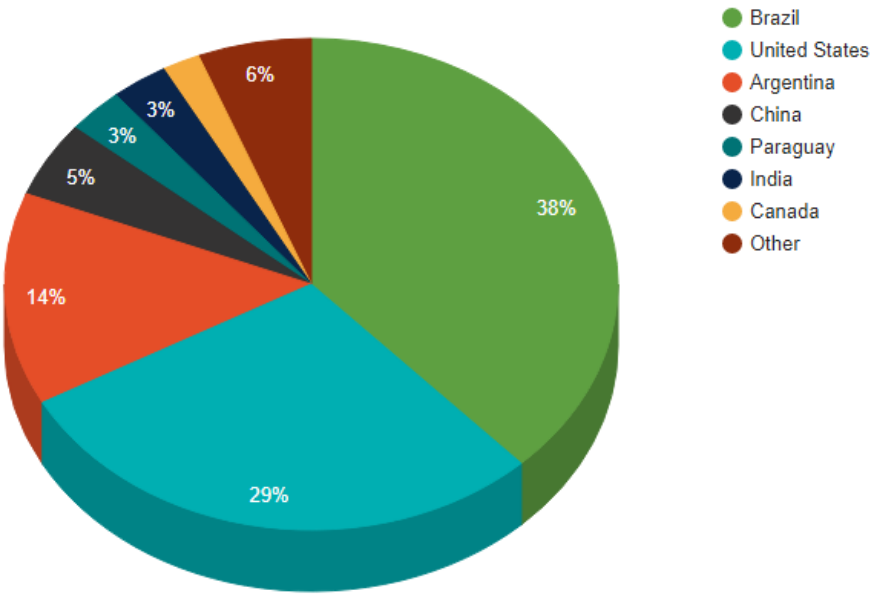

Figure 1.2: The Estimated World Soybean Production in 2021 (Source: Soystats, 2021)

1.2.2. South African Soybean industry

Soybean production is essential for the increasing demand in the human population worldwide and the growing industries for livestock (mainly pig and poultry) and the cattle industry (DAFF, 2010). In SA, the production of soybean is still on the rise and has greatly increased for more than 20 years. The local soybean production reached approximately 1.6 million tons in 2018, which is the largest production compared to the recent years (Fig. 1.3). In the 2020/2021 growing season, soybean production is estimated at 1.4 million tons

(Statistica, 2021). The leading producers by provinces in the 2019/2020 growing season are; Free State contributing 504 thousand metric tons (MT), followed by Mpumalanga with 429 thousand MT, and KwaZulu Natal contributing 105 thousand MT (Fig. 1.4). Soybean is currently produced in almost all the provinces in the country except for the Western Cape province, according to the estimates recorded by Statistica in 2021. In Africa, SA is currently the leading importer of soybean oilcake compared to other countries (Grain SA, 2016). In the upcoming years, an increase in soybean production will lead to SA not requiring any imports but being able to increase its exports.

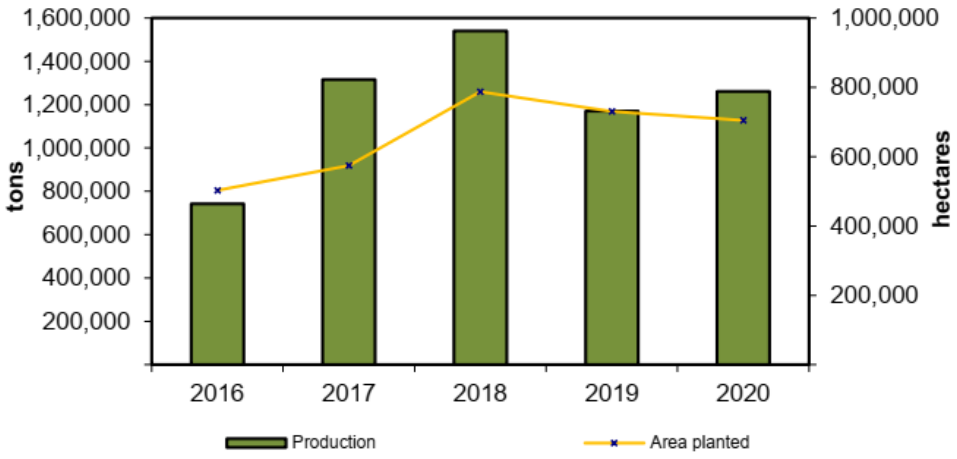

Figure 1.3: The Estimated Area Planted and Production of Soybean in SA from 2016-2020 (DAFF, 2019)

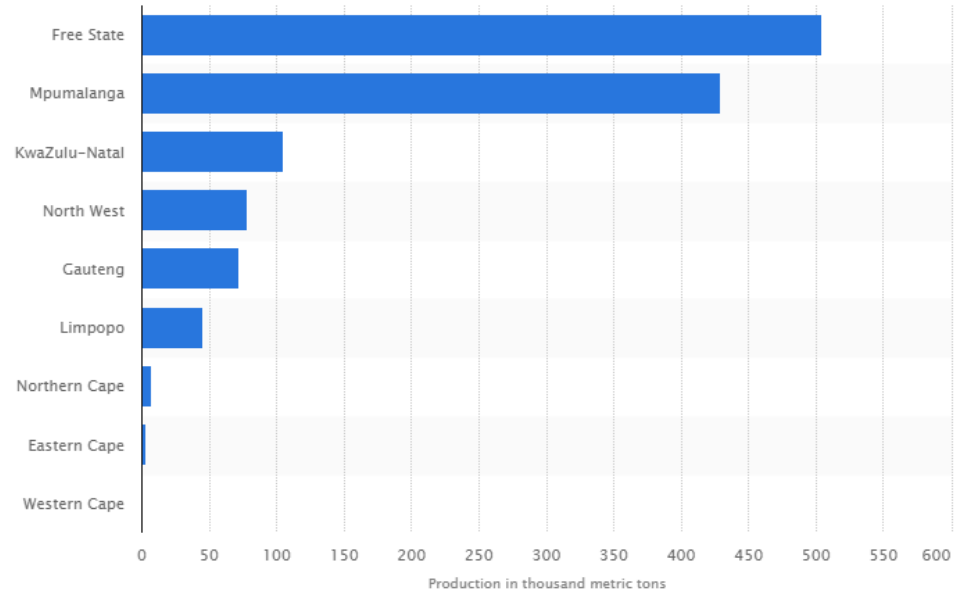

Figure 1.4: Soybean Distribution and Production in SA by Provinces (Statistica, 2021)

The local production of soybean is largely influenced by the global market. Although soybean production in SA had varied over time, a significant increase was observed in recent years. According to the Bureau for Food and Agricultural Policy (BFAP), 2019, the local soybean industry is faced with a shortage of new cultivars which has hindered its competitiveness. In addition, Africa has a challenge in the employment of modern agricultural methods and this results in slower improvements. However, SA is adopting the technology package used by the world's largest soybean producers (DAFF, 2017). New soybean cultivars are being introduced including PAN 1479R cultivar package launched in the 2020/2021 growing season (PANNAR, 2021). Research on plant diseases caused by fungi, bacteria and viruses has been done and more research is in progress.

1.3. The Current Status of Soybean Viruses in South Africa

A number of soybeans infecting viruses were reported worldwide and are estimated to be over 111 (Good and Monis, 2001). They belong to different virus genera and all these viruses infect soybean under natural conditions (Golnaraghi et al., 2004). These include economically important soybean viruses worldwide: soybean mosaic virus (SMV), bean pod mottle virus (BPMV), soybean vein necrosis virus (SVNV), alfalfa mosaic virus (AMV), peanut stunt virus (PSV), tobacco ringspot virus (TRSV), peanut mottle virus (PeMoV) and soybean dwarf virus (SbDV) (Hill and Whitham, 2014). Wherever soybean is grown, these viruses are likely to occur and reduce soybean yields. In SA, soybean was newly introduced in 1903 and farmers had limited information about the crop. However, in the study done by Pietersen et al., 1998 on the relative abundance of soybean viruses in South Africa, results showed the presence of several viruses. The losses incurred on soybean production by viruses was estimated at 20%. Since these studies were conducted several years ago, there is a possibility that new viruses could have emerged in the intervening years and those present could have evolved into new strains both contributing to significant yield losses. According to the latest International Committee on Taxonomy of Viruses (ICTV 2020, the viruses present in SA are classified under the following genera:

1.3.1. *Potyvirus*

Potyviruses belong to the family *Potyviridae*, which consists of the largest group of RNA plant virus species (ICTV, 2020). In South Africa, there are three potyviruses that were previously identified in soybean (Pieterson et al., 1998). These include soybean mosaic virus (SMV), peanut mottle virus (PeMoV), and cowpea aphid-borne mosaic virus (CABMV) according to the study done by Pietersen and Garnett, 1990; Pieterson et al., 1998. However, Klesser was the first to identify SMV in South African soybean in 1961. SMV is the most damaging and common virus disease in soybean worldwide, with infection resulting in severe mosaic and mottle symptoms (Surekha et al., 2018). There are risks of synergism due to

multiple infections; when SMV and bean pod mottle virus (BPMV) co-infect a soybean plant it results in severe symptoms (Hill et al., 2007; Nam et al., 2013). According to the research done in South Africa, BPMV has not been reported in soybean.

SMV is the most prevalent soybean virus worldwide including South Africa (Hajimorad et al., 2018). In a study done by Hajimorad et al., (2018), results showed a wide host range of SMV which included plants from the Solanaceae, Cucurbitaceae, Leguminosae and Caricaceae families. Symptoms produced include mottling, crinkling of leaves, leaf puckering, dwarfing and necrosis (Fig. 1.6). It comprises of a positive single stranded RNA genome encapsidated in flexuous, filamentous particles and range from 700-900nm in size (Zamora et al, 2017). The +ssRNA genome encodes for a single polyprotein open reading frame is translated into 10 proteins: P1- helper component proteinase, HC-Pro- helper component protease, P3- protein 3, 6K1- the 6 kDa protein 1, CI- cylindrical inclusion protein, 6K2- the 6 kDa protein 2, Nia- nuclear inclusion protein which is then processed into VPg- viral genome linked protein and NIa-Pro- NIa proteinase, Nib- nuclear inclusion protein b, CP- coat protein (Fig. 1.5). SMV is transmitted through soybean seed and by aphids in a non-persistent manner (Wang and Ghabrial, 2002).

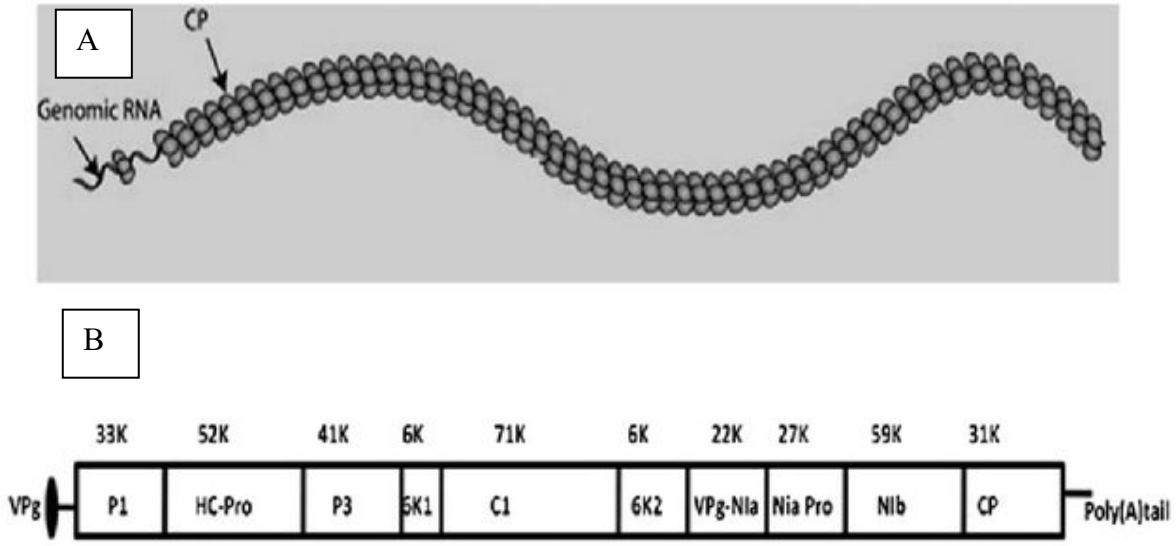

Figure 1.5: Genome Organisation and virus particle morphology for potyviruses. A. a flexuous rod-shaped *potyvirus* particle; B. A positive single stranded RNA genome for potyviruses encoding for 10 proteins (Mishra et al., 2013).

Figure 1.6: Bright yellow mosaic symptoms of SMV on soybean leaves under field conditions (Surekha et al., 2018).

1.3.2. *Alfamovirus*

Alfamovirus is a genus under the *Bromoviridae* family and consists of one species, alfalfa mosaic virus (AMV) according to the ICTV, 2020. The AMV genome is composed of a positive single stranded RNA tripartite genome with a size of 8.37kb (Bol, and Linthorst, 2003). The RNA genome is interestingly divided into four RNAs. RNA 1 and RNA 2 encode for the replicase proteins P1 and P2 respectively, RNA 3 encodes for the movement protein (P3) and Coat Protein (CP) (Fig. 1.7). Another CP is translated from the subgenomic RNA 4. The RNA, RNA 2 and RNA 3 are encapsidated separately into bacilliform virus particles with the size of 19 nm wide and 35-56 nm long (Xu and Nie, 2006; Bol and Linthorst, 2003). All the RNAs consist of a homologous sequence of 145 nucleotides.

AMV has a vast host range including the natural and experimental host range, with over 600 species from 70 families (Xu and Nie, 2006). The virus can be easily transmitted mechanically, through soybean seeds, by grafting and aphids in a non-persistent manner (He, 2011). On soybeans, the symptoms for AMV range from mosaic to mottle patterns (Fig. 1.8). AMV is a significantly important virus in soybean worldwide; wherever soybean is grown, it is likely to occur (Hill and Whitham, 2014). In previous studies on soybean viruses reported in South Africa, AMV was also identified (Pieterson et al., 1998).

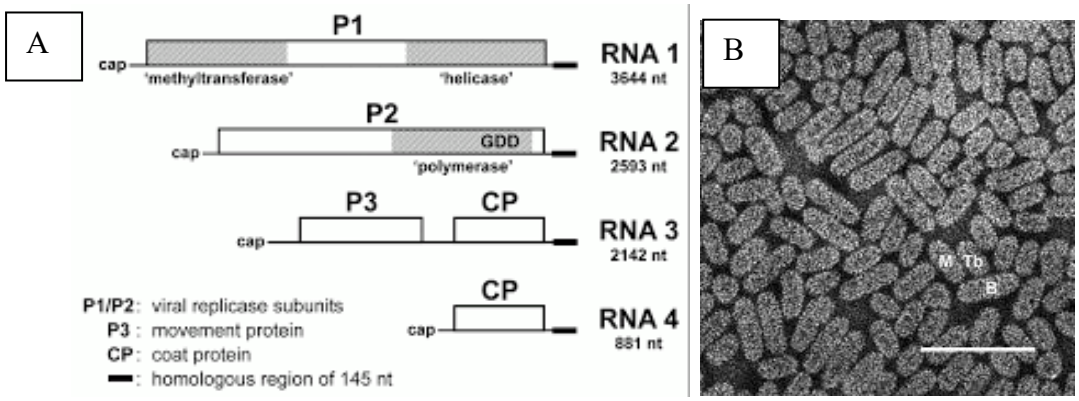

Figure 1.7: Genome organisation and particle morphology for AMV. A. negative single stranded RNA tripartite genome; B. Transmission Electron Micrograph of AMV shown by spheroidal and bacilliform virus particles (Bol and Linthorst, 2003)

Figure 1.8: Soybean leaves showing yellowing and mottling patterns of AMV (:Source: Crop Protection Network).

1.3.3. *Cytorhabdovirus*

Soybean blotchy mosaic virus (SbBMV) was newly identified in South Africa and found among the main soybean producing areas (Lamprecht et al., 2010). According to the ICTV, 2020, SbBMV is a related, unclassified virus species of cytorhabdoviruses. However, in the study done by Lamprecht et al. (2010), it was classified under the family *Rhabdoviridae* under the *Cytorhabdovirus* genus. Previously in a study done by Pietersen in 1990, a rhabdovirus associated with a blotchy mosaic symptom was identified. The symptoms for SbBMV include blotchy mosaic-like symptoms (Fig 1.10). Various leafhopper species are responsible for transmission. SbBMV has a potential threat to all the soybean crops grown in South Africa. However, to better understand this virus more studies still need to be conducted.

Cytorhabdoviruses generally encodes for 5 proteins with an addition of 1 protein for other species in the genus (Yang, 2017; ICTV, 2020). The RNA genome for SbBMV is translated to a total of 6 proteins (Fig. 1.9). N is the nucleocapsid protein gene responsible for encapsulating the viral genomic RNA to generate N–RNA complexes; two genes, 4a and 4b codes for the proposed phosphoprotein P and a putative protein respectively; M is the proposed matrix protein gene plays roles in the condensation of RNP complexes into a skeleton-like structure (RNP-M core) during virion assembly; G is the proposed glycoprotein gene forms transmembrane spikes (Walker et al., 2011). Additionally, the number of proteins may differ between cytorhabdoviruses, by encoding one or more accessory proteins between P and M and/or G and L coding regions, and L which is the proposed transcriptase gene (Walker et al., 2011).

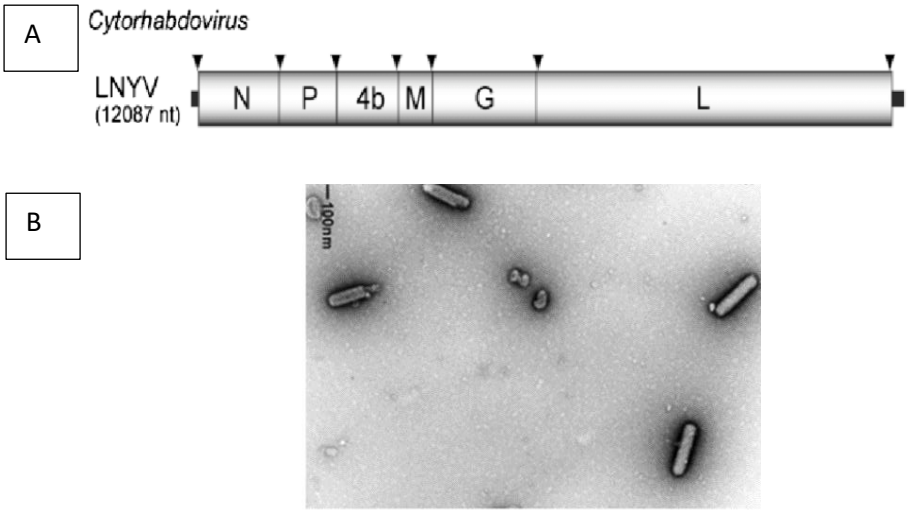

Figure 1.9: Genome organisation and virus particles for cytorhabdoviruses. A. Typical RNA genome for cytorhabdoviruses encoding for 6 proteins; B. A transmission electron micrograph for Soybean Blotchy Mosaic showing bacilliform shaped virions (Kondo, 2006; Lamprecht et al., 2010)

Figure 1.10: A soybean leaf showing blotchy mosaic symptoms caused by SbBMV (Strydom and Pietersen, 2018)

1.3.4. *Orthotospovirus*

Groundnut ringspot virus (GRSV) belonging to the family *Tospoviridae*, under genus *Orthotospovirus*, was reported in South African soybeans (Pieterson et al., 1998; Kuhn et al., 2020). However, the incidence was relatively low or not present in all the provinces (Pieterson et al., 1998). Orthospoviruses are characterised by spherical enveloped virus particles comprising a single stranded tripartite genome (Meng et al., 2015). The L (Large) RNA encodes for RNA-dependent RNA polymerase (RdRP), M (Medium) RNA encodes for the precursor of two viral glycoproteins (Gn/Gc) and a non-structural protein (NSm) which is involved in viral cell-to-cell movement and S (Small) RNA encodes for another non-structural protein (NSs) with RNA silencing suppression activity and the nucleocapsid (N) protein (Fig. 1.11).

According to a study done by Pietersen and Morris (2002), this was the first report of GRSV natural occurrence on soybean worldwide. Further analysis showed two new South African GRSV isolates present in the soybean plants (Fig. 1.12). Although studies were conducted, limited information is available on the distribution, incidence, and severity of this virus in SA. GRSV is transmitted by several species of Thrips (*Thripidae*) in a circulatory propagative manner (Lagos-Kutz et al., 2019.). The host range for GRSV is tomato, pepper, peanut, soybean, and coriander according to studies conducted in Brazil, Argentina, and SA (Leão et al., 2014). Symptoms include ringspots and necrotic, deformation of leaves and chlorotic ringspot on leaves and fruits on tomato and pepper (Golnaraghi et al., 2018; Webster et al., 2015).

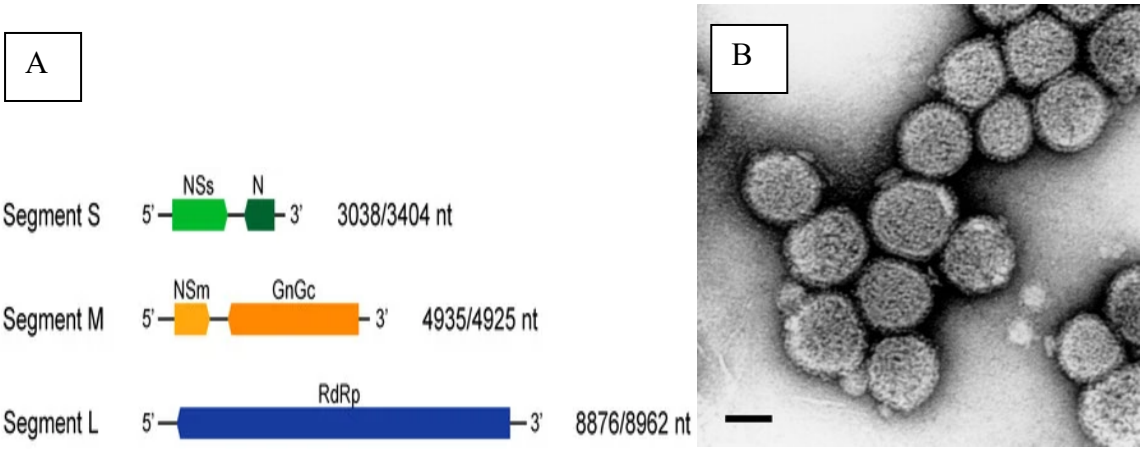

Figure 1.11: Genome organisation and particle morphology for Orthotospoviruses. A. Tripartite genome which encodes for 5 proteins; the transmission electron micrograph showing enveloped Spherical enveloped virus particles for Orthotospoviruses (Silva et al., 2019; Goldbach and Kormelink, 2011)

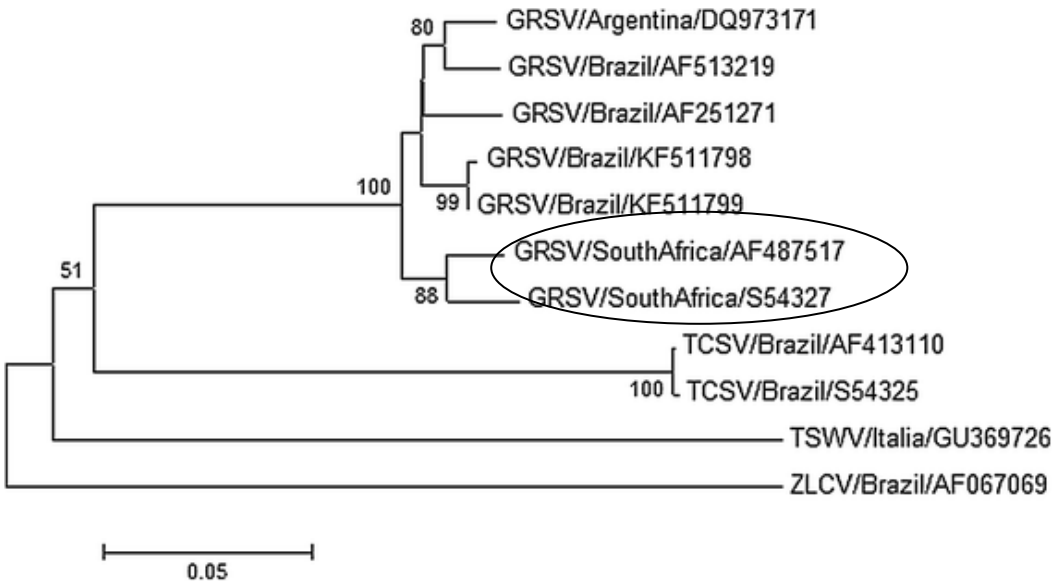

Figure 1.12: The Phylogenetic analysis of two South African GRSV isolates (Gürcan et al., 2021; Pietersen and Morris, 2002)

1.4. Identification of Soybean Viruses

1.4.1. Detection Methods for Plant Viruses

Virus spread from one crop to another is influenced mainly by climate change since it contributes to the evolution of viruses and their vectors' movement (Roth, 2020; McMichael et al., 2008). Plant virus symptoms are easily confused with physiological or nutritional deficiencies. On the other hand, different plant viruses cause similar symptoms and make it difficult to use symptoms for diagnosis. However, the most reliable way to confirm the presence of the viruses is by using virological techniques (Rubio et al., 2020). Many new technologies for virus identification have been developed over the years. At present, the most used detection methods for identification and characterisation of viruses include electron microscopy, enzyme-linked immunosorbent assay (ELISA), polymerase chain reaction (PCR), and Next-Generation Sequencing (NGS).

1.4.1.1. Polymerase Chain Reaction (PCR)

PCR is the most powerful tool used for virus identification ever since its discovery in the 1980s by Kary Mullis (Jeong et al., 2014). This method accurately targets a specific DNA sequence of the virus and amplifies it (Jeong et al., 2014). PCR utilizes the DNA polymerase to synthesize new DNA strands to make billions of copies exponentially (Santos et al., 2004). A programme is designed on a thermal cycler to perform PCR and the amplification is divided into three key steps; the first step is Denaturation, by heating at 90°C – 98°C; the second step is annealing, by cooling at 40°C – 65°C for the primers to bind the start and end of the targeted DNA; The last step is elongation by heating at 70°C – 75°C, the synthesis of the new DNA

strands starts from the primers (Garibyan and Avashia, 2013; Rychlik et al., 1991; Santos et al., 2004). The steps are repeated respectively for 20 – 40 cycles for the new DNA strand to be synthesized entirely (Garibyan and Avashia, 2013). For RNA viruses like Potyviruses, the RNA template is first converted to complementary DNA (cDNA) by Reverse Transcription PCR (Fig. 1.13).

PCR products are visualized by using preferably 1% or 1.5% of the gel electrophoresis to confirm the presents of a virus. PCR is the most convenient detection method due to time – it takes about 4 to 8 hours to obtain reliable results. However, PCR is very sensitive; any form of contamination can cause an error in the results. PCR products can be further characterized instantly by Sanger sequencing, enabling researchers to distinguish the different virus strains. More precise identification of virus species or isolates is achieved by comparing known sequences from the GenBank database.

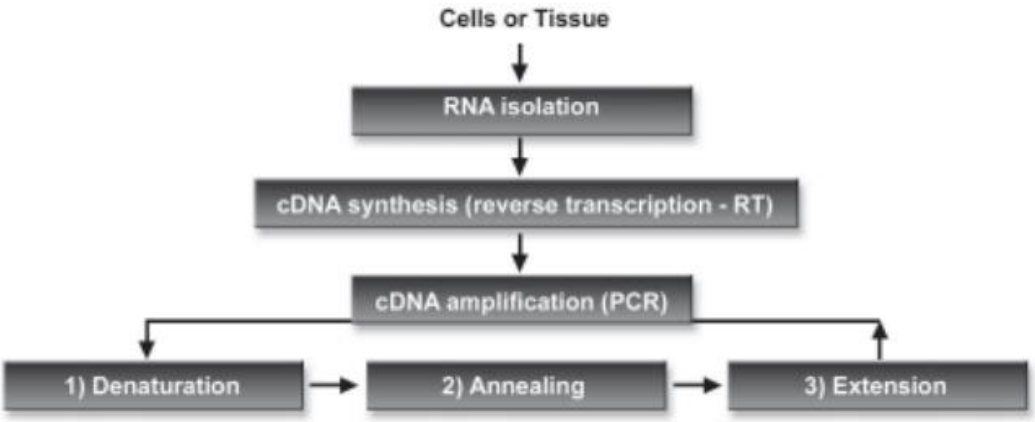

Figure 1.13: A schematic diagram showing RT-PCR and PCR used for detecting RNA viruses (Santos et al., 2004).

1.4.1.2. Enzyme-linked immunosorbent Assay (ELISA)

ELISA is a serological detection method based on the interaction of an antibody and antigen, with the addition of a substrate that generates a colour (Alemu, 2015). This method was introduced by Clark and Adams in 1977, has since been used successfully in detecting the virus's entire presence in plant tissue or cell (Boonham et al., 2014). Several types of ELISA have been developed over the years and these mainly include sandwich, competitive, direct, and indirect ELISA methods (Aydin, 2015). The sandwich ELISA is the most recommended compared to the other types because of its highest sensitivity (Aydin, 2015).

Triple antibody sandwich enzyme-linked immunosorbent assay (TAS-ELISA) and double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) are the two commonly utilized sandwich ELISAs for the detection of plant viruses (Seepiban et al., 2017).

In DAS-ELISA, two matched pair of antibodies are around the antigen and tagged to an enzyme that reacts with the substrate (Fig. 1.14). P-Nitrophenyl-phosphate (pNPP) is mostly used as the substrate and gives a colour for a positive reaction (Lorenz, 2011). The ELISA test is conducted in a 96-well microtiter plate preferably with rounded shape bottom, coated to bind protein very strongly (Auld et al., 2004). When low titre of antiserum is used viruses maybe be difficult to diagnose (Eick et al., 2016). However, ELISA was proven to be a sensitive, reliable, and cost-effective assay (Pandey et al., 2019; Aydin, 2015). Commercially available ELISA kits can be used for the detection and identification of soybean viruses.

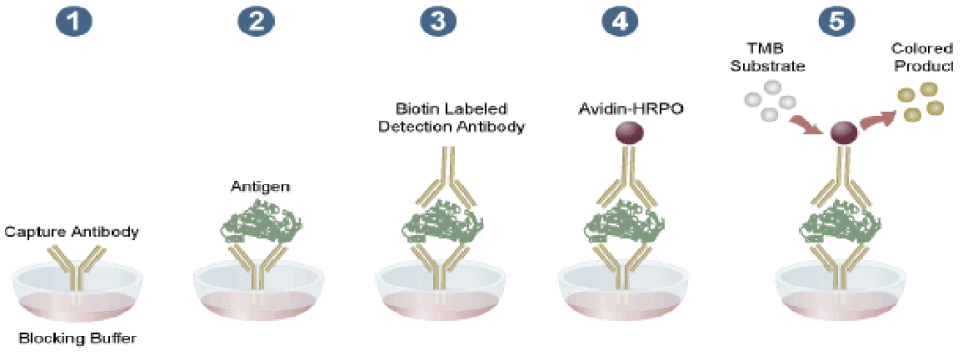

Figure 1.14: A schematic diagram showing the ELISA procedure from step 1 to step 5 (Alemu, 2015).

1.4.1.3. Next-Generation Sequencing (NGS)

NGS generates millions of sequenced data by fragmenting the extracted DNA or RNA into numerous copies in a parallel manner, subsequently adding specialized adapters, sequencing the libraries, and reassembling them to form a genomic sequence (Fig. 1.15). The applications of NGS in Plant Virology include RNA sequencing (RNA seq) and whole genome sequencing (WGS) (Barzon et al., 2011; Gupta and Verma, 2019). In SA, the NGS analysis can be done by Inqaba Biotech, ARC Biotechnology Platform, and Krisp Sanger Sequencing. The main types of sequencing technologies used are Illimina, Pac Bio, Life technologies (Ion Torrent, 454 Roche and SOLiD) and Nanopore. The level of accuracy is important when choosing these platforms for the analysis, the recommended platforms are Illumina also called Solexa Genome Analyser, 454 Roche, and SOLiD with 99,9% accuracy (Liu et al., 2012). These sequencing technologies generate raw data presented as either Fastq or BAM files with short read lengths (Frampton and Houlston, 2012). The read lengths differ for each platform, ~ 200 bp for Illumina/Solexa, ~700 bp for 454 Roche and ~85 bp for SOLiD (Liu et al., 2012).

Using this technique helps to identify the entire virus population present in the sample and detecting viruses even at low titers (Henson, Tischler, and Ning, 2012). NGS analysis has proven to be superior to the standard bioassay in detecting viruses of agronomic significance

(Boonham et al., 2014; Al Rwahnih et al., 2015). This may include its ability to discover unknown and known viruses and sequencing their full genome (Barzon et al., 2011; Rott et al., 2017; Chou, 2017). However, the interpretation of NGS data analysis is quite a challenging task without the background knowledge in Bioinformatics. Few bioinformatics pipelines are conveniently used, and they include DNASTAR which includes informative tutorials and Genome Detective Virus Tool. Further sequence analysis requires BLAST databases which is used to compare the virus genomes. These referencing databases are freely available on the Internet: http://www.uniprot.org/ and https://www.ncbi.nlm.nih.gov/.

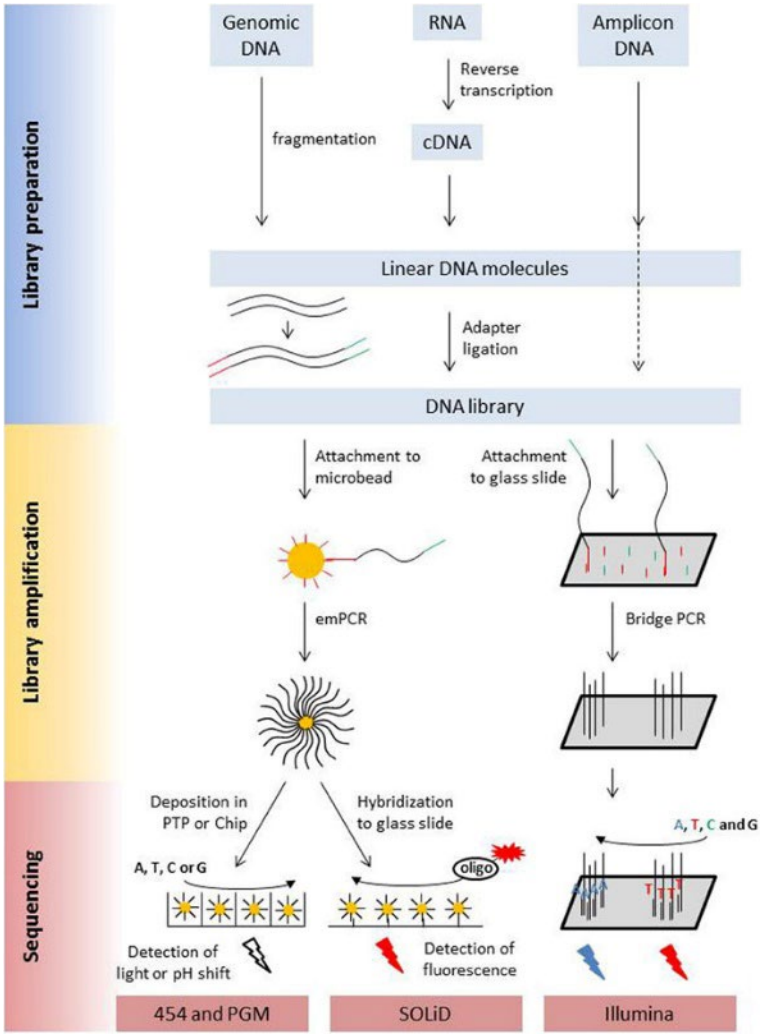

Figure 1.15: A schematic diagram illustrating the general NGS steps from library preparation to sequencing and commonly used platforms (Liu et al., 2016)

1.4.1.4. Electron Microscope (EM)

The EM is a standard tool used to examine the shape and size of the virus particles or virions (Milne and Vegetale, 2006; Zechmann and Zellnig, 2009). Electron Microscopy succeeds alone in the identification of viruses with distinct particles, like Rhabdoviruses and viruses from the *Closterovidae* family (Fig. 1.16). Rhabdoviruses are characterized by

bacilliform-shaped particles, which are approximately 60 × 120 nm in size (Liu et al., 2019). Closteroviruses have the longest flexuous rod-shaped particles, showing up to 2000nm in length when viewed under the Transmission Electron Microscope (German-Retana et al., 1999**)**. This technique has been very useful for diagnostics ever since it was invented in the early 1930s by Max Knoll and Ernst Ruska (Roingeard et al., 2019). TEM was frequently used to identify and characterise soybean viruses previously reported in South Africa. However, using both TEM and scanning electron microscope (SEM) helps to improve the characterisation of viruses with larger virions like baculovirus occlusion bodies (Gencer et al., 2018; Lopes et al., 2020).

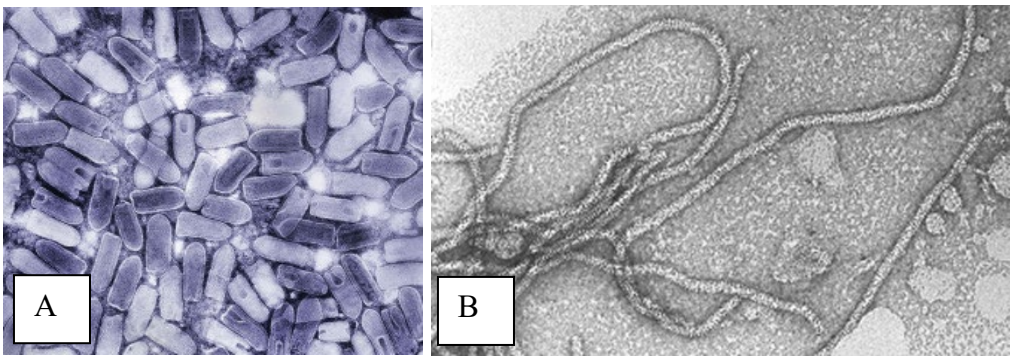

Figure 1.16: Electron micrographs showing A. Bacilliform particles for Rhabdoviruses; and B. Distinct flexuous rod-shaped *Closterovirus* particles observed by TEM with X25000 magnification (Goldblatt, 2013; German-Retana et al., 1999)

1.5. Control Methods for Soybean Viruses

In general, all viruses have no antiviral compounds available to cure plants infected with virus diseases. Accurate methods of virus detection and identification, both in plants and vectors, plays a pivotal role as a preliminary step for applying appropriate control measures. There are several approaches that can be used to limit the spread of viruses. Soybean viruses can be controlled using the following control methods:

1.5.1. Chemical control

Insecticides and pesticides have been utilised as chemical control to eradicate insect infestation on various crops and contributeto limiting the spread of soybean viruses (Hill and Whitham, 2014). However, there are side effects with insecticides, which can kill non-target species, contaminate the environment, and leave dangerous residues (Rocha and Villalobos, 2012). Seed treatments with a systemic insecticide are used for protecting young soybean plants, but the protection typically does not continue once the soybean plants reach maturity (Dashti et al., 2016). Agrochemicals are used as seed treatments, but little information is available about their efficacy against viruses. Many economically important soybean viruses

are vectored by whiteflies, aphids, and thrips. When the insect vector population is high, the viruses are transmitted and cause significant crop losses. The control of whiteflies using chemicals has not given any excellent result to decrease insect population (Murgianto and Hidayat, 2017). On the other hand, Thrips are not effectively controlled using insecticides due to their mobility, feeding behaviour and their protected egg and pupal stages (Bethke, Dreistadt, and Varela, 2014). Time, adequate spray coverage and targeting proper plants parts are key factors in the application of insecticides on thrips (Bethke, Dreistadt, and Varela, 2014). Soybean farmers must use more effective insecticides and use them in rotation or combination with appropriate biological and cultural practices to improve control (DAFF, 2010).

1.5.2. Biological control

This control method is commonly used and serves as alternative for controlling vectors that are responsible for transmitting viruses from one plant to another. It involves the preventative release of predators to regulate insect populations. The natural enemies of soybean aphids (*Aphis glycines*) are the ladybird beetles used as biological control (Hesler, 2014; Patterson et al., 2016). Infestations by aphids can affect soybean biomass and yield with the transmission of SMV, a devastating virus soybean virus (Revers and García, 2015). Lacewings, big-eyed bugs, and minute pirate bugs are all-natural enemies for whiteflies can be used as biological control (Flint, 1998). Predatory mites, minute pirate bugs, green lacewigs and a particular parasitic wasps feed on thrips and can be used as biological control (Loomans et al., 1995; Messelink et al., 2012). Farmers are recommended to use predatory mites and predatory bugs which are considered the most successful biological control agents against thrips, whiteflies, and aphids (Dreistadt, 2014). These predators colonize crops when the insect vectors are either absent or present at low densities because they can feed on alternative food sources (Dreistadt, 2014).

1.5.3. Cultural control

Commercial soybean cultivars are susceptible to SMV, the predominant devastating virus in soybean (Domier et al., 2007; Chawla et al., 2013; Mansky, 1990). Resistance genes were previously identified, and none of these genes are effective against all strains of the virus (Domier et al., 2007; Anderson, 2012). Resistance in soybean cultivars does not last long since viruses evolve due to mutations and recombination and break resistance. Studies have shown that RNA viruses (like SMV) mutate faster than DNA viruses forming new severe virus strains. Limited information is available and more resistant soybean cultivars still need to be investigated for use by farmers, so they can avoid the use of susceptible cultivars. Using virus-free seeds is the most important cultural control for many field crops, including soybean plants

(Akem, 1996; Nicaise, 2014; Roth et al., 2020). Integrated control is highly recommended for a farmer to best control soybean viruses.

1.6. Future Directions

The most incredibly challenging goal in agriculture is producing nutritious crops that are resistant to diseases and pests for the growing world population in a sustainable manner. According to the Food and Agriculture Organization of the United Nations (FAO) 2019, the world's population is expected to grow to approximately 10 billion by 2050. As a result, there is pressure to increase crop production in many countries. Brazil, Argentina, the United States of America, and China have invested more in research to increase yields and remain globally competitive in the soybean industry (FAO, 2017)

Limited Studies have been done in SA to identify the viruses infecting soybean and most of the studies were done many years ago. This is a challenge since viruses are one of the major constraints to soybean production. Emerging and undetected viruses pose a major challenge to soybean production. Virus identification plays a significant role as a preliminary first step in devising appropriate measures for virus disease control. NGS, PCR, ELISA and Electron Microscopy are considered the most accurate and reliable techniques used to identify and characterise plant viruses. Despite the growing cultivation of soybean across SA, limited research information is currently available on the present soybean viruses. More work is needed in the identification and characterization of viruses infecting soybean as well as developing appropriate control measures, which undoubtedly can help to improve soybean yields.

1.7. References

Alemu, K. (2015). Detection of diseases, identification, and diversity of viruses: A review. Journal of Biology, Agriculture and Healthcare, 5(1), 204-213.

Al Rwahnih, M., Daubert, S., Golino, D., Islas, C., and Rowhani, A. (2015). Comparison of Next-Generation Sequencing Versus Biological Indexing for the Optimal Detection of Viral Pathogens in Grapevine. Phytopathology, 105(6), 758–763.

Akem, C.N. (1996). Management of soybean diseases. Training Program, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. 36 p. Third edition.

Anderson, J.A. (2012). Evaluation of soybean recombinant inbred lines for yield potential and resistance to Sudden Death Syndrome. Carbondale: Southern Illinois University.

Auld, D. S., Coassin, P. A., Coussens, N. P., Hensley, P., Klumpp-Thomas, C., Michael, S., Sittampalam, G.S., Trask, O.J., Wagner, B.K., Weidner, J.R., Wildey, M.J. and Dahlin, J.L.,

(2004). Microplate Selection and Recommended Practices in High-throughput Screening and Quantitative Biology, Assay Guidance Manual, (Md), 1–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/32520474.

Aydin, S. (2015). A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA, Peptides. Elsevier Inc., 72, 4–15.

Barzon, L., Lavezzo, E., Militello, V., Toppo, S., and Palù, G. (2011). Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. International Journal of Molecular Sciences, 12(11), 7861–7884.

Bethke, J., Dreistadt, S. and Varela, L. (2014). 'Tree squirrels - integrated pest management for home gardeners and landscape professionals', Pest Notes. University of California. Agriculture and Natural Resources, 1–8.

BFAP (2019). The Bureau for Food and Agricultural Policy. https://www.bfap.co.za/wp-content/uploads/2020/04/Final-Baseline-2019.pdf. Accessed: 03 August 2021.

Bol, J. F., and Linthorst, H. J. M. (2003). Alfalfa mosaic virus: coat protein-dependent initiation of infection. Molecular Plant Pathology, 4(1), 1–8.

Boonham, N., Kreuze, J., Winterc, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., Mumford, R. (2014). Methods in virus diagnostics: From ELISA to next-generation sequencing. Virus Research 186:20-31.

Chawla, S., Bowen, C. R., Slaminko, T. L., Hobbs, H. A., and Hartman, G. L. (2013). A public program to evaluate commercial soybean cultivars for pathogen and pest resistance. Plant Disease, 97(5), 568–578.

Chou, W. C., Lin, S. S., Yeh, S. D., Li, S. L., Peng, Y. C., Fan, Y. H., and Chen, T. C. (2017). Characterization of the genome of a phylogenetically distinct *tospovirus* and its interactions with the local lesion-induced host Chenopodium quinoa by whole-transcriptome analyses. PLoS ONE, 12(8), 1–23.

Crop Protection Network 2016. Alfalfa mosaic of soybean. Available at the website: https://cropprotectionnetwork.org/resources/articles/diseases/alfalfa-mosaic-of-soybean. Accessed: 25 July 2021

DAFF (2019). Department of Agriculture, Fisheries and Forestry. Report 2019.

DAFF (2017). Department of Agriculture, Fisheries and Forestry. Report 2017.

Dashti, N. H., Cherian, V. M., Smith, D. L., and Mcgill, J. (2016). Soybean production. Abiotic and Biotic Stresses in Soybean Production, 217–240.

Devine, T.E, and Card, A. (2013). Forage soybeans. In: Rubiales D, editor. Legume Perspectives: Soybean: A Dawn to the Legume World.

Dlamini, T. S., Tshabalala, P., and Mutengwa, T. (2014). Soybeans production in South Africa. OCL - Oilseeds and Fats, 21(2).

Domier, L. L., Steinlage, T. A., Hobbs, H. A., Wang, Y., Herrera-Rodriguez, G., Haudenshield, J. S., and Hartman, G. L., 2007. Similarities in seed and aphid transmission among Soybean mosaic virus isolates. Plant Disease, 91(5), 546–550.

Dreistadt, S. H. (2014). 'Biological control and natural enemies of invertebrates', University of California Agriculture and Natural Resources Publication 74140. 1–6.

Du Toit, J.J. (1942). The Cultivation of soybeans. Farming in South Africa.

Eick, G., Urlacher, S. S., McDade, T. W., Kowal, P., and Snodgrass, J. J. (2016). Validation of an Optimized ELISA for Quantitative Assessment of Epstein-Barr Virus Antibodies from Dried Blood Spots. Biodemography and Social Biology, 62(2), 222–233.

Engelbrecht, G., Claassens, S., Mienie, C. M. S., and Fourie, H. (2020). South africa: An important soybean producer in sub-saharan africa and the quest for managing nematode pests of the crop. Agriculture (Switzerland), 10(6), 1–19.

Figueiredo, P. N. (2016). New challenges for public research organisations in agricultural innovation in developing economies: Evidence from Embrapa in Brazil's soybean industry. The Quarterly Review of Economics and Finance, 62, 21–32.

FAO (2019). Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data/QC (verified June 2021). Accessed: 29 July 2021

FAO (2017). Food and Agriculture Organization of the United Nations. The future of food and agriculture – Trends and challenges. Rome.

Flint, M. L. (1998). Pests of the Garden and Small Farm. 2nd ed. Oakland University of California, Berkeley. Division of Agriculture and Natural Resources. 3332.21

Frampton, M., and Houlston, R. (2012). Generation of Artificial FASTQ Files to Evaluate the Performance of Next-Generation Sequencing Pipelines. PLoS ONE, 7(11), e49110.

Garibyan, L. and Avashia, N. (2013). Polymerase chain reaction, Journal of Investigative Dermatology. Elsevier Masson SAS, 133(3), 1–4.

German-Retana, S., Candresse, T., and Martelli, G. (1999). Closteroviruses (*Closteroviridae*). Encyclopedia of Virology, 266–273.

Goldbach, R., and Kormelink, R. (2011). Tospovirus. The Springer Index of Viruses, 231–235.

Goldblatt, A. (2013). 'Agriculture: Facts and Trends South Africa 1'. Available at: http://awsassets.wwf.org.za/downloads/facts_brochure_mockup_04_b.pdf. Accessed: 11 Novemeber 2020.

Golnaraghi, A. R., Shaheen, N., Pourrahim, R., Farzadfar, S., and Ghasemi, A. (2004). Occurrence and the relative incidence of viruses infecting soybeans in Iran. Plant Disease, 88(10), 1069–1074.

Golnaraghi, A., Shahraeen, N. and Nguyen, H. D. (2018). Characterization and Genetic Structure of a Tospovirus Causing Chlorotic Ring Spots and Chlorosis Disease on Peanut; Comparison with Iranian and Polish Populations of Tomato yellow fruit ring virus, Plant Disease, 102(8), 1509–1519.

Good, X. and Monis, J. (2001). Partial genome organization, identification of the coat protein gene, and detection of Grapevine leafroll-associated virus-5, Phytopathology, 91(3), 274–281.

Gürcan, K., Turan, S., Teber, S., Kılınçer, İ., Uz, İ., Tamisier, L., and Çağlayan, K. (2021). Molecular and biological characterization of a new mulberry idaeovirus. Virus Research, 298, 198411.

Guo J., Wang Y., Song C., Zhou J., Qiu L., Huang H., and Wang Y. (2010). A single origin and moderate bottleneck during domestication of soybean (Glycine max): implications from microsatellites and nucleotide sequences. Annals of Botany, 106(3), 505-514.

Gupta N., and Verma V.K. (2019). Next-Generation Sequencing and Its Application: Empowering in Public Health Beyond Reality. In: Arora P. (eds) Microbial Technology for the Welfare of Society. Microorganisms for Sustainability, Springer, Singapore, 17, 313-341.

Grain SA, 2016. South Africa's soybean industry: A brief overview. Available at https://www.grainsa.co.za/south-africa-s-soybean-industry:-a-brief-overview. Accessed: 01 August 2021

Hajimorad, M. R., Domier, L. L., Tolin, S. A., Whitham, S. A., and Saghai Maroof, M. A. (2018). Soybean mosaic virus: a successful potyvirus with a wide distribution but restricted natural host range. Molecular Plant Pathology, 19(7), 1563–1579.

Hartman GL, West E.D., and Herman TK. (2011). Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Security 3(1):5-17.

He, B., Hill, J. H. and Hajimorad, M. R. (2011). Factors to Improve Detection of Alfalfa mosaic virus in Soybean, Plant Health Progress, 12(1), 8.

Hesler, L. S. (2014). Inventory and Assessment of Foliar Natural Enemies of the Soybean Aphid (Hemiptera: Aphididae) in South Dakota. Environmental Entomology, 43(3), 577–588.

Henson, J., Tischler, G., and Ning, Z. (2012). Next generation sequencing and large genome assemblies. Pharmacogenomics, 13(8), 901–915.

Hill, J. H., Koval, N. C., Gaska, J. M., and Grau, C. R. (2007). Identification of field tolerance to Bean pod mottle virus and Soybean mosaic virus in soybean. Crop Science. 47:212.

Hill, J. H., and Whitham, S.A. (2014). Control of Virus Diseases in Soybeans. Advances in Virus Research, 90, 355-390.

ICTV (2020). International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/ Accessed: 04 June 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.

Kim, M.Y., Lee, S., Van, K., Kim, T.H., Jeong, S.C., Choi, I.Y., Kim, D.S., Lee, Y.S., Park, D., Ma, J. et al. (2010). Whole-genome sequencing and intensive analysis of the undomesticated soybean (Glycine soja Sieb. and Zucc.) genome. Proceedings of the National Academy of Sciences, 107(51), 22032-22037.

Kim, M. Y., Van, K., Kang, Y. J., Kim, K. H., and Lee, S. H. (2011). Tracing soybean domestication history: From nucleotide to genome. Breeding Science, 61(5), 445–452.

Kofsky, J., Zhang, H., and Song, B. H. (2018). The untapped genetic reservoir: The past, current, and future applications of the wild soybean (*Glycine soja*). Frontiers in Plant Sciences, 9, 949.

Kondo, H. (2006). Orchid fleck virus is a rhabdovirus with an unusual bipartite genome. Journal of General Virology, 87(8), 2413–2421.

Kuhn, J. H., Adkins, S., Alioto, D., Alkhovsky, S. V., Amarasinghe, G. K., Anthony, S. J., and Zhou, X. (2020). 2020 taxonomic update for phylum Negarnaviricota (Riboviria: Orthornavirae), including the large orders Bunyavirales and Mononegavirales. Archives of Virology (Vol. 165), 3023–3072.

Lamprecht, R. L., Kasdorf, G. G. F., Stiller, M., Staples, S. M., Nel, L. H., and Pietersen, G. (2010). Soybean blotchy mosaic virus, a new Cytorhabdovirus found in South Africa. Plant Disease, 94(11), 1348–1354.

Lagos-Kutz, D., Pawlowski, M. L., Haudenshield, J., Han, J., Domier, L. L., and Hartman, G. L. (2019). Evaluation of Soybean for Resistance to Neohyadatothrips variabilis (Thysanoptera: *Thripidae*) Noninfected and Infected with Soybean Vein Necrosis Virus. Journal of Economic Entomology, 113(2), 949–955.

Leão, E. U., Spadotti, D. M. de A., Rocha, K. C. G., Pantoja, K. de F. da C., Rezende, J. A. M., Pavan, M. A., and Krause-Sakate, R. (2014). *Citrullus lanatusis* a New Natural Host of Groundnut ringspot virus in Brazil. Journal of Phytopathology, 163(11-12), 1014–1018.

Liu, K. S. (2008). Food Use of Whole Soybeans. Soybeans: Chemistry, Production, Processing, and Utilization, 441–481.

Liu, J. Z., Fang, Y., and Pang, H. (2016). The current status of the Soybean-Soybean Mosaic Virus (SMV) pathosystem. Frontiers in Microbiology, 7(11), 1–9.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., and Law, M. (2012). Comparison of Next-Generation Sequencing Systems. Journal of Biomedicine and Biotechnology, 2012, 1–11.

Li, Y., Zhao, S., Ma, J., Li, D., Yan, L., Li, J., and Qiu, L. (2013). Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. BMC Genomics, 14(1), 579.

Liu, W., Fan, Y., Li, Z., Zhao, J., Zhou, Y., Jiang, N., Zeng, J., Cain, K., and Zeng, L. (2019). Isolation, identification, and classification of a novel rhabdovirus from diseased Chinese rice-field eels (Monopterus albus). Archives of Virology, 164(1), 105–116.

Loomans, A. J. M. Lenteren, J. C. V., Tommasini, M. G., Maini, S., and Riudavets, J. (1995). Biological control of thrips pests, Wageningen Agricultural University Papers. Available at: http://www.cabdirect.org/abstracts/19961100108.html. Accessed: 15 October 2020

Lopes, W., Cruz, G. N. F., Rodrigues, M. L., Vainstein, M. H., Kmetzsch, L., Staats, C. C., and Schrank, A. (2020). Scanning electron microscopy and machine learning reveal heterogeneity in capsular morphotypes of the human pathogen *Cryptococcus* spp. Scientific Reports, 10(1).

Lorenz, U. (2011). Protein tyrosine phosphatase assays. Current Protocols in Immunology, 93, 11.7.

Makurira, H. (2010). Rainfed Agriculture in Sub-Saharan Africa. Water Productivity in Rainfed Agriculture, 181, 9–21.

Mansky, L. M. (1990). Interactions of soybean mosaic virus with susceptible and resistant lines of soybean. Iowa State University, Ames, IA 50011, United States.

Messelink, G. J., Bloemhard, C. M. J., Sabelis, M. W., and Janssen, A. (2012). Biological control of aphids in the presence of thrips and their enemies. BioControl, 58(1), 45–55.

McMichael, A. J., and Woodruff, R. E. (2008). Climate change and infectious diseases. The Social Ecology of Infectious Diseases, 378–407.

Meng, J., Liu, P., Zhu, L., Zou, C., Li, J., and Chen, B. (2015). Complete Genome Sequence of Mulberry Vein Banding Associated Virus, a New Tospovirus Infecting Mulberry. PLoS ONE, 10(8), e0136196.

Milne, R. G. and Vegetale, V. (2006). Electron Microscopy as a Powerful Tool for Detection and Identification of Plant Viruses. In XI International Symposium on Virus Diseases of Ornamental Plants 722, 37–40.

Mishra, R., Verma, R. K., Sharma, P., Choudhary, D. K., and Gaur, R. K. (2013). Interaction between viral proteins with the transmission of Potyvirus. Archives of Phytopathology and Plant Protection, 47(2), 240–253.

Murgianto, F., and Hidayat, P. (2017). Whitefly Infestation and Economic Comparison of Two Different Pest Control Methods on Soybean Production. Planta Tropika: Journal of Agro Science, 5(2).

Mullis, K. B, (1990). The unusual origin of the polymerase chain reaction. Scientific American, 262(4), 56–65.

Nam, M., Bae, H., Hammond, J., Domier, L. L., Youn, Y.-N., Lee, B.-C., and Lim, H.S. (2013). Seed Transmission Rates of Bean pod mottle virus and Soybean mosaic virus in Soybean May Be Affected by Mixed Infection or Expression of the Kunitz Trypsin Inhibitor. Research in Plant Disease, 19(2), 114–117.

Nicaise, V. (2014). Crop immunity against viruses: outcomes and future challenges. Frontiers in Plant Science, 5, 660.

OECD (2001). Organisation for Economic Co-Operation and Development. Adoption of Technologies for Sustainable Farming Systems Wageningen Workshop Proceedings. Available at: https://www.oecd.org/greengrowth/sustainable-agriculture/2739771.pdf. Accessed: 01 June 2021.

Otulak, K., and Garbaczewska, G. (2012). Cytopathological Potato virus Y structures during Solanaceous plants infection. Micron, 43(7), 839–850.

Pandey, A. K., Varshney, R. K., Sudini, H. K., and Pandey, M. K. (2019). An Improved Enzyme-Linked Immunosorbent Assay (ELISA) Based Protocol Using Seeds for Detection of

Five Major Peanut Allergens Ara h 1, Ara h 2, Ara h 3, Ara h 6, and Ara h 8. Frontiers in nutrition, 6, 68.

Patterson, R., Extension, C. C., Ramirez, R., and Specialist, E. E. (2016). Aphid Natural Enemies and Biological Control.

Pietersen, G., and Garnett, H. M. (1990). Survey for the viruses of soybeans (Glycine max. L) in the Transvaal, South Africa. Phytophylactica, 22, 35-40.

Pietersen, G., Staples, S. M., Kasdorf, G. G. F., and Jooste, A. E. C. (1998). The relative abundance of soybean viruses in South Africa. African Plant Protection, 4(2), 65–70.

Pietersen, G., and Morris, J. (2002). Natural Occurrence of Groundnut ringspot virus on Soybean in South Africa. Plant Disease, 86(11), 1271–1271.

Revers, F., and García, J. A. (2015). Molecular biology of potyviruses. Advances in virus research, 92, 101–199.

Rocha, P. J. and Villalobos, V. M. (2012). Comparative study of genetically modified and conventional soybean cultivation in Argentina, Brazil, Paraguay and Uruguay Inter-American Institute for Cooperation on Agriculture. ISBN: 978-92-9248-444-6

Roingeard, P., Raynal, P. I., Eymieux, S., and Blanchard, E. (2019). Virus detection by transmission electron microscopy: Still useful for diagnosis and a plus for biosafety. Reviews in Medical Virology, 29(1), e2019.

Roth, M. G., Webster, R. W., Mueller, D. S., Chilvers, M. I., Faske, T. R., Mathew, F. M., and Smith, D. L. (2020). Integrated Management of Important Soybean Pathogens of the United States in Changing Climate. Journal of Integrated Pest Management, 11(1), 17.

Rott, M., Xiang, Y., Boyes, I., Belton, M., Saeed, H., Kesanakurti, P., and Rast, H. (2017). Application of Next Generation Sequencing for Diagnostic Testing of Tree Fruit Viruses and Viroids. Plant Disease, 101(8), 1489–1499.

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(6), 1–23.

Rychlik, W., Spencer, W. J. and Rhoads, R. E. (1991). Optimization of the annealing temperature for DNA amplification in vitro, Nucleic Acids Research, 19(3), 698.

Santos, C. F. dos, Sakai, V. T., Machado, M. A. de A. M., Schippers, D. N., and Greene, A. S. (2004). Reverse transcription and polymerase chain reaction: principles and applications in dentistry. Journal of Applied Oral Science, 12(1), 1–11.

Sedivy, E. J., Wu, F., and Hanzawa, Y. (2017). Soybean domestication: the origin, genetic architecture and molecular bases. New Phytologist, *214*(2), 539–553.

Seepiban, C., Charoenvilaisiri, S., Warin, N., Bhunchoth, A., Phironrit, N., Phuangrat, B., and Gajanandana, O. (2017). Development and application of triple antibody sandwich enzyme-linked immunosorbent assays for begomovirus detection using monoclonal antibodies against Tomato yellow leaf curl Thailand virus. Virology Journal, 14(1), 99.

Shea, Z., Singer, W. M. , and Zhang, B. (2020). Soybean Production, Versatility, and Improvement. In (Ed.), Legume Crops - Prospects, Production and Uses. IntechOpen.

Singh, P., Kumar, R., Sabapathy, S. N., and Bawa, A. S. (2008). Functional and edible uses of soy protein products. Comprehensive Reviews in Food Science and Food Safety, 7(1), 14–28.

Silva, J., de Oliveira, A., de Almeida, M., Kormelink, R., Nagata, T., and Resende, R. (2019). Tomato Chlorotic Spot Virus (TCSV) Putatively Incorporated a Genomic Segment of Groundnut Ringspot Virus (GRSV) Upon a Reassortment Event. Viruses, 11(2), 187.

SoyStats (2021). A reference guide to important soybean facts and figures. American Soybean Association. http://soystats.com/ Accessed 05 June 2021.

Surekha, S., Magar, S. J., Satyadev P. and Prasanna Krishna V. (2018). Host Range of Aphid Transmissible Mosaic Virus Infecting Cowpea. International Journal of Current Microbiology and Applied Sciences. 7(6): 3080-3085.

Statistica (2021). Total number of processed soya beans in South Africa from 2000 to 2021, by type (in metric tons). https://www.statista.com/statistics/1139019/soya-beans-production-in-south-africa-by-province/ Accessed: 05 June 2021.

Strydom, E. and Pietersen, G. (2018). 'Alternative hosts and seed transmissibility of soybean blotchy mosaic virus'. European Journal of Plant Pathology 151 (1), 263-268.

Szostak, B., Głowacka, A., Kasiczak, A., Kiełtyka-Dadasiewicz, A., and Bąkowski, M. (2020). Nutritional value of soybeans and the yield of protein and fat depending on a cultivar and the level of nitrogen application. Journal of Elementology, 25(1), 45–57.

The National Center for Biotechnology Information (NCBI), 1998. https://www.ncbi.nlm.nih.gov/ Accessed: 01 October 2021.

The Universal Protein Resource (UniProt), 2008. https://www.uniprot.org/ Accessed: 28 July 2021.

USDA (2021). World agricultural supply and demand estimates, the report for August 2021. Available at: http://www.usda.gov/oce/commodity/wasde/. Accessed: 16 September 2021.

Voora, V., Larrea, C. and Bermudez, S. (2020). 'Global Market Report: Soybeans'. Available at https://www.iisd.org/system/files/2020-10/ssi-global-market-report-soybean.pdf. Accessed: 09 June 2021.

Wang, R. Y., and Ghabrial, S. A. (2002). Effect of Aphid Behavior on Efficiency of Transmission of Soybean mosaic virus by the Soybean-Colonizing Aphid (*Aphis glycines*). Plant Disease, 86(11), 1260–1264.

Walker, P. J., Dietzgen, R. G., Joubert, D. A., and Blasdell, K. R. (2011). Rhabdovirus accessory genes. Virus Research, 162(1-2), 110–125.

Webster, C. G., Frantz, G., Reitz, S. R., Funderburk, J. E., Mellinger, H. C., and McAvoy, E. (2015). The emergence of Groundnut ringspot virus and Tomato chlorotic spot virus in vegetables in Florida and the south-eastern United States. Phytopathology, 105, 388-398.

Xu, H. and Nie, J. (2006). Identification, characterization, and molecular detection of Alfalfa mosaic virus in Potato, Phytopathology, 96(11), 1237–1242.

Yang, X., Huang, J., Liu, C., Chen, B., Zhang, T., and Zhou, G. (2017). Rice Stripe Mosaic Virus, a Novel Cytorhabdovirus Infecting Rice via Leafhopper Transmission. Frontiers in microbiology, 7, 2140.

Zamora, M., Méndez-López, E., Agirrezabala, X., Cuesta, R., Lavín, J. L., Amelia Sánchez-Pina, M., and Valle, M. (2017). Potyvirus virion structure shows conserved protein fold and RNA binding site in ssRNA viruses. Science Advances, 3(9), 1–8.

Zechmann, B., and Zellnig, G. (2009). Rapid diagnosis of plant virus diseases by transmission electron microscopy. Journal of Virological Methods, 162(1-2), 163–169.

Chapter 2

Detection and Identification of Viruses Infecting Soybean (*Glycine max*. L) In KwaZulu Natal Province, South Africa

Abstract

Soybean (*Glycine max* L.) is the world's most important seed legume, mainly used as an oil crop and protein source. Many destructive plant viruses have been found to infect soybean on a global scale. In South Africa, few viruses have been detected infecting soybean crops. The studies of viruses on soybean were undertaken several years ago and it is possible that the virus population structure may have evolved over time. The aim of this study was to detect and identify viruses presently infecting soybean grown in the province of KwaZulu-Natal, South Africa and determine the incidence of any seed transmitted viruses. Fifty-four soybean leaf samples exhibiting virus-like symptoms were collected from breeding lines growing in a Plant Pathology disease garden and greenhouses at the University of KwaZulu-Natal during the 2018 - 2019 and 2019 - 2020 growing seasons. Mechanical inoculation using inoculum prepared from the soybean field samples was done on *Nicotiana tabacum* L. to propagate the viruses in the collected samples. Symptom development was monitored on inoculated *N. tabacum* plants for 2-3 weeks after inoculation. The field samples were also subjected to Reverse Transcription Polymerase Chain Reaction (RT-PCR) and PCR to detect viruses known to infect soybean worldwide. Generic and specific primers were used to target specific coding regions of the viruses tested. Antibodies specific to cucumber mosaic virus (CMV) and tobacco mosaic virus (TMV) were used to test for virus presence in the field samples using double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA). The study also focused on determining the incidence of seed transmitted viruses by planting different soybean varieties/lines and testing for virus presence after the plants had germinated. Virus presence was based on symptoms exhibited by the germinated plants and by DAS-ELISA. Results of this study showed that the inoculated *N. tabacum* plants developed virus-like symptoms. Soybean mosaic virus (SMV), TMV, CMV, and hibiscus chlorotic ringspot virus (HCRSV) were identified in the field samples based on PCR results. Seed transmission assays did not demonstrate the presence of viruses based on symptomatology and DAS-ELISA tests. In conclusion, our results showed that the population structure of viruses infecting soybean in South Africa has changed since the last studies were done. This is the first report of CMV, TMV and HCRSV infecting soybean in South Africa.

2.1. Introduction

Soybean (*Glycine max* L.) is the most significant seed legume in the world, primarily used as an oil crop and a protein source (Pietersen and Garnett, 1990; Golnaraghi et al., 2004). In South Africa (SA), soybean production was estimated at 1.4 million tonnes in 2020 (Statistica, 2021). KwaZulu-Natal province is among the largest soybean producers in South Africa, with an estimate of 105 thousand metric tonnes produced in 2020 (Statistica, 2021). Plant viruses are a major threat to soybean production worldwide. Currently, more than 111 viruses reported worldwide are known to infect soybean (Hill and Whitham, 2014). Economic losses from soybean due to virus infection are estimated at several billion US dollars (Hartman et al., 2011). Soybean mosaic virus (SMV) is the most devastating seed transmitted virus infecting soybean (Golnaraghi et al., 2004). The most economically important viruses infecting soybean worldwide include SMV, bean pod mottle virus (BPMV), tobacco ringspot virus (TRV), soybean vein necrosis virus (SVNV), soybean dwarf virus (SDV), peanut mottle virus (PeMoV), peanut stunt virus (PSV) and alfalfa mosaic virus (AMV) (Hill and Whitman, 2014).

In SA, the following viruses have been reported to infect soybean; SMV, cowpea aphid-borne mosaic virus (CABMV), alfalfa mosaic virus (AMV), peanut mottle virus (PeMoV), and groundnut ringspot virus (GRSV) (Pietersen et al., 1998; Pietersen and Morris, 2002; Pietersen and Garnet, 2000). Some of these viruses are known to be seed transmitted (He et al., 2011; Wang and Ghabrial, 2002). In a study conducted by Lamprecht et al., (2010), a new virus belonging to the *Rhabdoviridae* family, soybean blotchy mosaic virus (SbBMV), was also identified infecting soybean. According to our knowledge, limited studies have been conducted on viruses infecting soybean in South Africa in the last 10 years. Viruses evolve, so there is every likelihood that there has been a change in the status of viruses occurring in South Africa since the last study was conducted.

Plant viruses are currently detected using a variety of virological techniques. Polymerase Chain Reaction (PCR) and double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) are the most common due to their high sensitivity (Rahbari et al., 2021; Aydin, 2015). However, according to Jeong et al. (2014), PCR is the most sensitive detection method compared to ELISA in plant virus detection. It is very crucial to detect viruses early and accurately in order to develop suitable and effective virus disease control measures. Against this background, this study aimed to detect and identify viruses presently infecting soybean in KwaZulu-Natal, South Africa. In addition, the incidence of any seed transmitted viruses in a number of soybean cultivars/lines was determine

2.2. Materials and methods

2.2.1. Sample Collection

Soybean breeding lines exhibiting virus-like symptoms were collected from the Plant Pathology disease garden and greenhouses at the University of KwaZulu-Natal during the 2018 - 2019 and 2019 - 2020 growing seasons. A total of 54 symptomatic soybean leaves showing bright yellow mosaic, chlorotic spots, mottle, chlorotic spots, stunted growth, leaf malformation, and vein clearing (Fig. 2.1) were sampled. These symptoms were observed on different soybean varieties/breeding lines including Jacaranda, Mukwa, 1454 R, 1521 R, and an unknown line. Soybean leaves exhibiting the symptoms described above were harvested and packaged into clearly labelled plastic bags and stored at -80℃ for longterm until analysis.

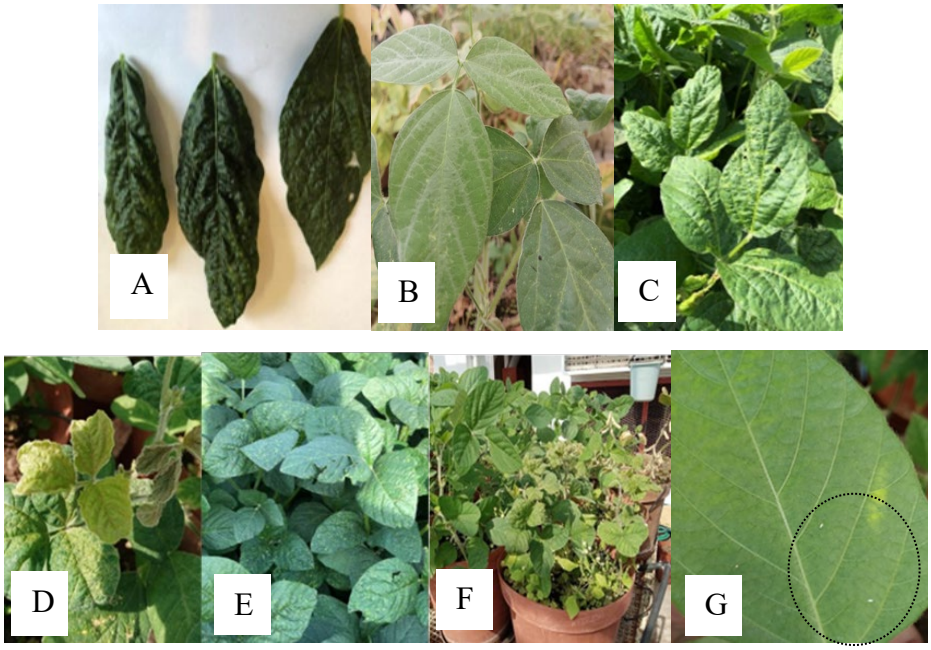

Figure 2 . 1: Soybean samples exhibiting virus-like symptoms observed on different soybean lines growing in the Plant Pathology disease garden and greenhouses. (A) mottle, (B) Vein clearing, (C) Bright yellow mosaic, (D) Leaf malformation, (E) Chlorotic Spots, (F) Stunted growth, and (G) whiteflies observed on the soybean samples.

2.2.2. Virus Propagation

Nicotiana tabacum L. seeds were sown in 30cm pots containing compost potting media. Irrigation was done once a day by supplementing water with 3:1:3 (38) soluble NPK fertilizer and plants germinated after 2-3 weeks. Seedlings were transplanted to 20cm pots then grown until the 5-6 leaf stage to begin inoculation. Plants were maintained in an insect-proof greenhouse which was kept at a constant temperature of 25℃. Soybean samples exhibiting virus-like symptoms were used to prepare inoculum which was used to inoculate the *N. tabacum* plants to propagate the unknown virus present in the sample. The inoculum

was prepared by grinding soybean leaves using a mortar and pestle with the addition of liquid Nitrogen and 0.1M phosphate buffer (pH 7.0) which was maintained at 4℃. Carborundum abrasive powder was dusted on the *N. tabacum* leaves to enhance wounding of leaves during the inoculation process. Inoculation was achieved by gently rubbing (to avoid damaging the leaves) a pestle dipped in freshly prepared inoculum on leaves dusted with carborundum. Inoculated leaves were then rinsed with distilled water to heal the damaged leaf cells. Inoculated plants were monitored for any symptom development over a 2–3-week period after inoculation. Furthermore, symptomatic leaves were harvested and subjected to RT-PCR and PCR to confirm virus presence.

2.2.3. RNA and DNA Extraction

Total RNA and DNA was extracted from each of the soybean field leaf samples exhibiting virus-like symptoms using Quick RNA™ MiniPrep Kit (Zymo Research, USA), Quick RNA™ MicroPrep Kit (Zymo Research, USA) and Quick-DNA™ Microprep Kit according to the manufacturer's instructions provided in each kit. 0.5 g for each sample was homogenized in a sterile 1.5ml Eppendorf tube using sterile drill bits and liquid nitrogen. 600 µL of the lysis buffer and sixplastic beads were added to each tube then vortexed. Thereafter, they were centrifuged at 16,000 x g for 2 min in a (14000rpm) microcentrifuge (SIGMA, USA). 400 µL of the clear lysate from each tube was transferred in a 1.5ml sterile Eppendorf tube containing 400 µL of 95-100% ethanol. The solution was mixed thoroughly by pipetting and transferred into the Zymo-Spin™ column in a Collection Tube to be centrifuged at 16,000 x g for 30 sec. The flow-through collected by the collection tube was discarded. 400 µL of RNA Prep Buffer was added to the column and centrifuged for 30 sec. Followed by two washing steps; 700 µl RNA Wash Buffer added to the column, centrifuged for 30 sec and discarded flow-through, and then added 400 µL RNA Wash Buffer and centrifuge the column for 2 min and discarded flow-through. The column was transferred in a 1.5ml sterile Eppendorf tube to elute 15 µL DNase/RNase-Free Water directly to the column matrix, centrifuged for 30 sec. The RNAs were immediately stored at 4 ℃ to be analysed for quality and quantity using the Nanodrop ND 1000 Spectrophotometer (Inqaba Biotech, SA) and stored at -80 ℃ until further analysis.

2.2.4. RT-PCR and PCR

Soybean field leaf samples were tested for different viruses known to infect soybean using RT-PCR and PCR (Table 2.1). For RNA viruses, total RNA was extracted from the symptomatic soybean leaves was used as template for the RT-PCR. The pre-RT step was initially done to linearize RNA complex structure by incubating 4 µL of the RNA template at 65℃ for 5 minutes. The template was added in the RT-PCR master mix including 2 µL of 10µM reverse primer, 2 µL of dNTPs mix, 4 µL of 5X Reaction Buffer, 1 µL of Revert aid, 1 µL of RNAse inhibitor (Ribolock) used to prevent the RNA from degrading and 6 µL of Nuclease-

free water. These reagents were obtained from the Reverse Aid premium reverse transcriptase kits (Thermo Fisher Scientific Inc, USA). RT-PCR conditions were as follows; 20 µL of the total master mix for each sample were incubated at 42℃ for 60 min and heated at 70℃ for 10 min. For DNA viruses, DNA was extracted and directly subjected to PCR. In each PCR master mix, 4 µL of the DNA and complementary DNA (cDNA) obtained from RT-PCR was added including 2 µL for each of the Forward and reverse Primers (Inqaba biotech, SA) listed in Table 1, and 2 µL Nuclease-free water and 10 µL DreamTaq PCR Master Mix (2X) (Thermo Scientific, USA) for each sample.

PCR parameters for SMV, BCMV and BCMNV primers listed in Table 2.1 included 35 cycles; Denaturation step at 95℃ for 45 sec, annealed for 2 min according to the specific primer annealing temperature in Table 2.1 and the elongation step at 72℃ for 2 min and the extended step at 72℃ for 10 min. The PCR parameters for Carlavirus, cowpea mild mottle virus (CpMMV), nuclear inclusion b (NIB) protein for potyviruses, Begomovirus and bean pod mottle virus (BPMV) primers as listed in Table 2.1 included 40 cycles; Denaturation step was at 98℃ for 1 s, annealed for 20 sec as per primer in Table 1, and the elongation step at 72 ℃ for 20 sec. The PCR parameters for Tospovirus, TMV, CMV (polyvalent primers) and HCRSV primers were set for 35 cycles; Denaturation step was at 95℃ for 30 secannealed for 1min according to the specific primer annealing temperature in Table 2.1, and elongation was done at 72℃ for 1 min and 5 min for the extended elongation step. All steps from Pre-RT, RT-PCR and PCR were performed using the PCR G-Storm Thermal Cycler (Gene Technologies, UK).

2.2.5. Agarose Gel Electrophoresis

PCR products were examined on a 1.5% (w/v) agarose gel electrophoresis. For gel preparation, 1.05 g of the gel powder was poured in 250ml Erlenmeyer flask. It was then diluted with 70 ml of 1X Tris-acetate-EDTA (TAE) Buffer with an addition of 5ml to account for evaporation errors. Subsequently, the gel was heated on a microwave for 1 min to dissolve, then cooled for 5 min and stained with SYBR® safe stain (Invitrogen, USA). The gel was poured on a cast with a comb on one end and left to solidify for 10 min. Thereafter, PCR products were mixed with 3µL of 6X DNA Loading Dye (Thermo Scientific, USA) for each sample. A GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) and GeneRuler 1 kb DNA Ladder (Thermo Scientific, USA) were added on the gel with the PCR products to measure the DNA sizes. DNAs were visualized and captured in the G: BOX with GeneSnap (Syngene software version 7.12.). The electrophoresis power supply was set on 100 V and 400 mA for the duration of 45 min to run the gel.

2.2.6. ELISA

Symptomatic leaf tissue of 30 soybean breeding lines/varieties collected from soybean field samples and germinated seedlings including Jacaranda, Mukwa, 1454 R, 1521 R, 1623 R and an unknown line were tested using readily available monoclonal antibodies, specific for CMV and TMV by the standard double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA). This test was conducted to confirm results from PCR. Buffer preparations and all DAS-ELISA steps were followed according to manufacturers' instructions of DAS-ELISA diagnostic kit (SEDIAG, France). 100 µL of the Coating antibodies for CMV and TMV were diluted in 10ml coating buffer 1X pH 9.6. 100 µL was added in each well used and incubated at 37℃ for 2 hrs. The plate was washed threetimes in Microwash 1100 ELISA Plate Washer (Inqaba Biotech, SA) using the washing buffer. 1 g of each of the 10 soybean samples including the positive and negative controls was ground using mortar and pestle in 10ml of extraction buffer pH 7.3 [PBS 1X, PVP (Mw 10, 000-40,000) and Tween 20] for homogenization.

Mortar and pestle were rinsed thoroughly in between sample preparation to avoid contamination. 100 µL of sample was added to each well and plates were incubated at 4℃ for 16 hrs (overnight). The plate was then washed threetimes in Microwash 1100 Elisa Plate Washer (Inqaba Biotech, SA) using the washing buffer. 100 µL of the antibody-enzyme conjugate specific for CMV and TMV were diluted in 10ml of conjugate buffer 1X pH 7.3. 100 µL was then added to each well used and incubated at 37℃ for 2 hrs. This was followed by the washing step done threetimes to remove any unbound antibodies. Lastly, 100 µL of 1mg/ml of the enzyme substrate 4-Nitrophenyl phosphate disodium salt hexahydrate (pNPP) was added to each well and incubated at 37℃ for 2 hrs in the absence of light. The 96-well microplate was evaluated for any colour change as an indication of a positive reaction. No colour change was an indication of a negative reaction. Absorbance reads at 450nm is an indication of the antigen present. The 96-well microplate was also analysed by reading the microplate reader (Inqaba Biotech, SA). The absorbance readings were expected to be three times more compared to the negative control to be considered a positive reaction.

2.2.7. Seed Transmission Assays

Soybean varieties/lines including 1454 R, 1521 R, and 1623 R were provided by Pannar Seeds (Pty), Greytown, South Africato be used in this study. The seed-lots from each variety/line provided were too many to count, in each experiment 100 seeds were randomly selected from each variety to be planted in a 300-hole seedling tray containing compost potting media. This experiment was repeated fivetimes by growing seedlings using the same seed-lots from the three varieties. Irrigation was done daily, and seedlings were evaluated for virus-like symptom development after 2-3 weeks of germination. Seedlings were maintained at a

constant temperature of 25°C in a polycarbonate tunnel, an insect-proof greenhouse, to assess the incidence of seed transmission. Germinated plants were harvested and tested using DAS-ELISA following the procedure described in section 2.2.6. The virus presence was based on symptoms exhibited by the germinated plants and by the DAS-ELISA test.

2.3. Results

2.3.1. Virus Propagation

Two weeks after inoculation, typical virus-like symptoms were observed on the inoculated *N. tabacum* plants (Fig. 2.2). Leaves from symptomatic plants were harvested and packaged in a clearly labelled plastic bag and then stored at -80 ℃ for further analysis. RT-PCR results for the inoculated samples were negative for viruses tested.

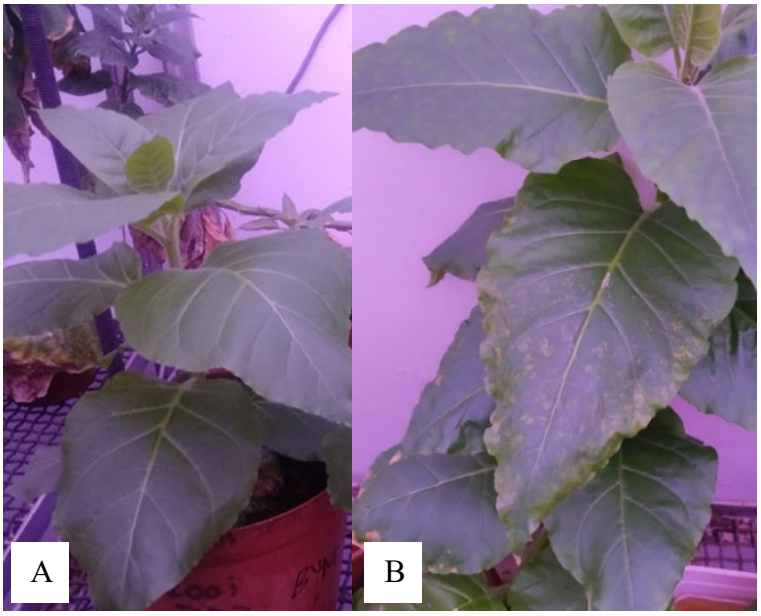

Figure 2.2: *N. tabacum* plants mechanically inoculated for propagating the viruses. (A) Healthy plant (control); and (B) Typical mosaic-like virus symptoms shown on tobacco plant inoculated using soybean plants exhibiting bright yellow mosaic and mottle symptoms.

2.3.2. RT-PCR

An amplicon of the expected size of ~ 469 bp was observed from 10of the 54 soybean symptomatic leaf samples tested which confirmed the presence of SMV (Fig. 2.3). An expected band size of 512 bp was observed on five of the samples tested which confirmed the presence of TMV (Fig. 2.4). Another expected band size of 380 bp was observed in four of the soybean samples tested) confirming the presence of CMV (Fig. 2.5). The presence of HCRSV was also confirmed in two of the soybean samples tested when an expected band size of 420 bp was observed (Fig. 2.6).

Table 2.1: List of Primer sets used for the amplification of all viruses tested in this study.

Primers	Sequences 5' – 3.'	Amplification Products	Annealing temperature	References
SMV-CP-F SMV- CP-R	TCAGGCAAGGAGAAGGAAGG CTGCGGTGGGCCCATGC	469 bp	55℃	Chen et al., 2015
BPMV: RNA2-fwd RNA2-rev	ATACCCCTAATGGCACAGGA GGAAATGTAACCACCCGAAT	132 bp	60℃	Pflieger et al., 2014
D1BCMNV RU	GAGGTGTATGAATCCGTGTCAAC TCAGTATTCTCGCTGGTTGTTGC	350bp	55.5℃	Chiquito-Almanza et al., 2017
D1BCMV RU	AAATGTGGTACAATGCTGTGAAC TCAGTATTCTCGCTGGTTGTTGC	350 bp	55.5℃	Chiquito-Almanza et al., 2017
Carla335R Carla322F	GCTAATTTCGTAGACCAGAGAG CTCGAGTACCTAGAGAGGAAGAG	131 bp	50℃	Li et al., 2013
Carla5468F Carla5672R	CCATGGCCAATATGCTTTTCAC CCATCAGCGCATAAGTTCCACC	~1500 bp	46℃	Li et al., 2013
Tospovirus: gM410 (F) gM870c(R)	AACTGGAAAAATGATTYNYTTGTTGG ATTAGYTTGCAKGCTTCAATN ARGC	500 bp	52°C	Chen et al., 2012
CpMMV-CPu CpMMV-CPd	TTTACTCTTAggTWATggAgTC CCTATTAAAACACACAAHTCAAA	867 bp	52℃	Chang et al., 2013
Begomovirus: Forward Reverse	GGRTTDGARGCATGHGTAC ATGGCCYATRTAYAGRAAGCCMAG	~550bp	51℃	Marwal et al., 2013
NIB 3F NIB 2R	GTITGYGTIGAYGAYTTYAAYAA TCIACIACIGTIGAIGGYTGNCC	350 bp	45℃	Zheng et al., 2010

TMV R TMV F	CGATGATGATTCGGAGGC GAGGTCCARACCAAMCCAG	512 bp	52℃	Kimaru et al., 2020
PfCMVRNA3-1163 PrCMVRNA3-2034	ATGCTTCTCCRCGAGATT GTAAGCTGGATGGACAAC	871 bp	55℃	Chang, McLaughlin and Tolin 2011
HC F8 HC R3	TGGGATGGAGGTGAAGCAGA AAGGGCTGCCTCACAACTATGG	420 bp	52℃	Niu et al., 2014; Zheng et al., 2018
HCRSV F HCRSV R	AAGAGAGCAGCCAATAGA GAAGAAGAACAAGAAGCGA	759 bp	57℃	Li and Chang, 2002
HCRSV Forward HCRSV Reverse	GGAACCCGTCCTGTTACTTC ATCACATCCACATCCCCTTC	557 bp	55℃	Shafie, 2019
Polyvalent Primers (CMV RNA 1) Bromo-Fw **(20µM)** Bromo-Rv **(20µM)** Bromo-RT Adaptor-RC5 Adaptor-RC3	**CCAACGGAATTCCTCACTAAAC**GCTCCYCAYR GICTKGCTGGTGCYCT **CACATCGGAACTCGGTACCT**CCCATRTCATAA CCICCATGIAT RTCRAACAICATIGCICCRTCGAACAT **CCAACGGAATTCCTCACTAAAC** **CACATCGGAACTCGGTACCT**	380 bp	52℃	Seo et al., 2014

Note: [a.] Polyvalent Primers (CMV RNA 1) show non-viral adaptor sequences represented in bold.

[b.] Concentration for all the primers was 10µM excluding Bromo-Fw and Bromo-Rv with 20µM as Indicated in bold.

[c.] Bromo-RT primer was used in RT-PCR step, then Bromo-Fw, Bromo-Rv, Adaptor RC 5 and RC 3 primers were used in the PCR step

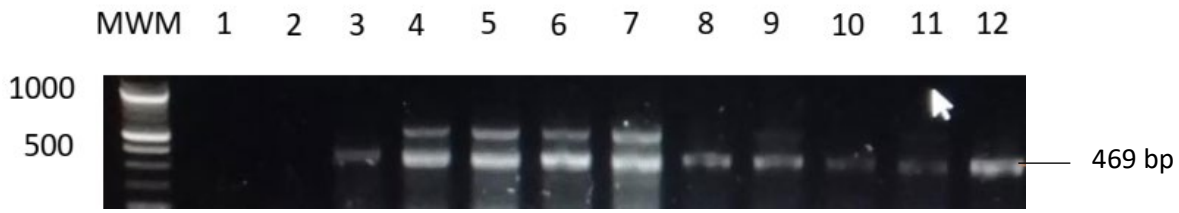

Figure 2.3: An agarose gel electrophoresis showing the RT-PCR products for the coat protein gene of SMV from the Plant Pathology disease garden soybean samples. The amplification size was obtained by using primers designed by Chen et al., 2015. This figure shows a GeneRuler 100 bp DNA Ladder (Thermo Scientific, USA) contained in lane 1. In lane 2 and 3, no band sizes were observed which indicates that SMV was not detected. Samples in lanes 4-13 have band sizes of ~ 469 bp showing that SMV might be present.

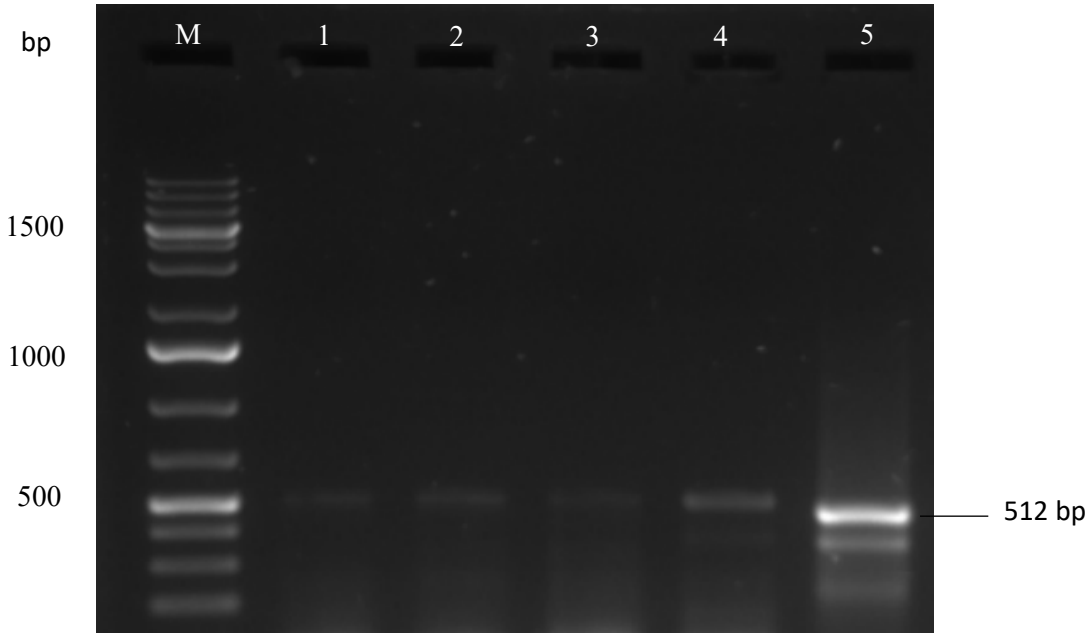

Figure 2.4: An agarose gel electrophoresis showing the RT-PCR products for the coat protein gene of TMV from the Plant Pathology disease garden soybean samples. The amplification size was obtained by using primers designed by Kimaru et al., 2020. This figure shows a GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) contained in lane 1. In lanes 1-5, a band size of 512 bp shows that TMV is present.

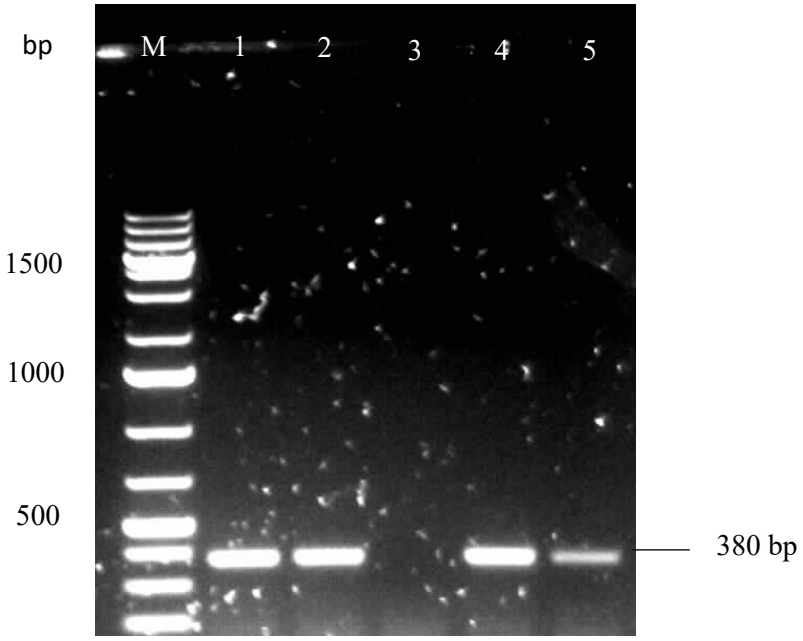

Figure 2.5: An agarose gel electrophoresis showing the RT-PCR products for the coding region Segment RNA 1 of CMV from the Plant Pathology disease garden and greenhouse soybean samples. The amplification size was obtained by using polyvalent primers designed by Seo et al., 2014. This figure shows a GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) contained in Lane 1. Lane 4 (soybean sample 3) showed that CMV was not detected. Samples in Lanes 2, 3, 5, and 6 have a band size of 380 bp show that CMV is present.

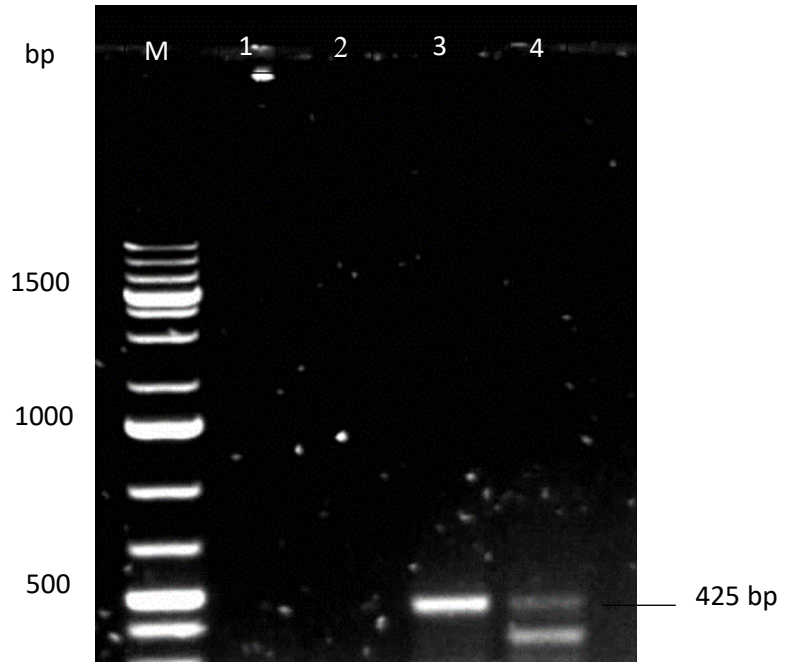

Figure 2.6: An agarose gel electrophoresis showing the RT-PCR products for the CP gene of HCRSV from greenhouse collected soybean samples. The amplification size was obtained

using primers designed by Niu et al., 2014. This figure shows a GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) contained in Lane 1. Lane 2 and 3 indicates 2 soybean samples tested negative for HCRSV. Samples in Lanes 4 and 5 have a band size of 425 bp indicates that HCRSV is present.

2.3.3. DAS-ELISA

All the soybean varieties/lines were negative for CMV and TMV after testing using DAS-ELISA (Table 2.2). There was no colour change observed on the 96-well microtiter plate except for the positive controls.

Table 2.2: Soybean varieties/lines used for the detection of CMV and TMV by DAS ELISA

Varieties/lines	Number of samples	DAS-ELISA Reaction
Jacaranda	3	–
Mukwa	2	–
1454 R	8	–
1521 R	8	–
1623R	8	–
Unknown	1	–

Note:

[a.] – is an indication of all the soybean samples that tested negative

2.3.4. Seed Transmission Assays

No virus-like symptoms were observed when seedlings of the varieties were monitored for any symptom development. All the three soybean varieties/lines did not show any symptom development 2-3 weeksafter germination. These results were concluded after the experiment was repeated fivetimes. All the germinated seedlings were negative for the presence of CMV and TMV when tested using DAS-ELISA.

2.4. Discussion

This study aimed to identify the viruses infecting some of soybean cultivars and breeding lines grown in the province of KwaZulu-Natal, South Africa. The approach used was to first propagate the virus on *N. tabacum* so as to have enough virus to work with. The results from this part of the study showed typical mosaic-like virus symptoms on the inoculated *N. tabacum* plants. However, the inoculated plants were negative when tested for viruses using RT-PCR. This may be due to the fact that the symptoms observed on the field were as a result of mixed virus infections and the symptoms observed on the inoculated plants might have been caused by other viruses in the mixture other than CMV and TMV that were tested for. It

is important to point out that the main focus of this study was to test all soybean field samples exhibiting virus-like symptoms for virus presence using PCR and ELISA.

Some viruses known to infect soybean worldwide and whitefly-transmitted viruses were tested on all the soybean field samples used in the study. Although whitefly vectors were observed on some of the soybean leaves sampled for testing, Whitefly-transmitted viruses including begomoviruses were not detected in the soybean field samples using PCR. The testing for a wider range of whitefly-transmitted viruses other than the ones tested for in this study would have yielded a different set of results. However, RT-PCR successfully detected four viruses which included SMV, TMV, CMV and HCRSV (Fig 3-6). These results showed that bright yellow mosaic symptoms on soybean samples were as a result of TMV and CMV infection while chlorotic spots were caused by CMV. Stunted growth, vein clearing, chlorotic spots and leaf malformation symptoms were caused by HCRSV infection. Interestingly, this is the first study to report HCRSV infecting soybean worldwide. According to a recent study done by Shafie (2019), HCRSV has a wide host range which include members of the families *Chenopodiaceae*, *Malvaceae*, *Solanaceae* and two members from the family *Fabaceae* (*Vigna unguiculata* L. Walp and *Phaseolus* vulguris L.. HCRSV symptoms on hibiscus plants include stunted growth, chlorotic or ring spots, mottling and vein banding (Ramos-González et al., 2020). Similarly, in this study, some of these symptoms were observed on the infected soybean plants. There is need to study HCRSV infecting soybean further.

DAS-ELISA was used to test for CMV and TMV in the infected field samples. Our results showed that all samples tested were negative for the two viruses (Table 2). This result is not surprising given that ELISA tests at times fail to detect the virus when it is present in low concentration in the tested sample (Erick et al., 2016). When the virus titre is low or the virus concentration is under the technique detection threshold, this can limit the sensitivity of the ELISA test leading to a negative result (Rubio et al., 2020).

Seed transmission occurs at a relatively very low rate when vector activity is controlled (Nam et al., 2013; Abney and Plopper, 1994). For seed-transmission of viruses to occur, an infected mother plant must pass the virus onto the offspring, which are the newly produced plants (Cobos et al., 2019). When an infected seed is planted, the virus may occur on the emerging seedling. The results from this study showed no incidences for seed transmission of viruses in the 1521 R, 1454 R and 1623 R soybean varieties based of symptoms on germinated seedlings and DAS-ELISA tests. The seeds used in the study might have been produced under optimal conditions which minimised vector activity and hence infection of the plants resulting in little or no seed transmission. The results from the seed transmission study are therefore not surprising.

2.5. Conclusion

Accurate detection and identification of viruses plays an important role in virus disease management. Results from this study showed that SMV, TMV, CMV, and HCRSV are the viruses infecting the soybean field samples that were evaluated. Molecular detection techniques were used to accurately detectand identify the viruses infecting soybean field samples. The findings from our study showed that the population structure of viruses infecting soybean in SA has changed over the years since studies on soybean viruses were last done. According to our knowledge to date, this is the first report of CMV, TMV and HCRSV infecting soybean in South Africa. This is also the first report of HCRSV infecting soybean worldwide.

This study serves as basis for the development strategies to effectively manage viruses infecting soybean in SA. Future studies should focus on using Next Generation Sequencing (NGS) to give a complete picture of the viruses infecting soybean in SA given that mixed virus infection on crops is very common. Using NGS will also allow the studying of full genome sequences of the viruses identified leading to a better understanding of their biology. It is important to emphasize that early and accurate detection of the viruses is very crucial for developing and then applying effective and sustainable control measures of the diseases they cause.

2.6. References

Abney, T. S., and Plopper, L. D. (1994). Effects of Bean pod mottle virus on soybean seed maturation and seed borne *Phomopsis* spp. Plant Disease, 78, 33-37.

Aydin, S. (2015). A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA, Peptides. Elsevier Inc., 72, 4–15.

Chang, C. A., Chien, L. Y., Tsai, C. F., Lin, Y. Y., and Cheng, Y. H. (2013). First Report of Cowpea mild mottle virus in Cowpea and French bean in Taiwan. Plant disease, 97(7), 1001.

Chang, P. G. S., McLaughlin, W. A., and Tolin, S. A. (2011). Tissue blot immunoassay and direct RT-PCR of cucumoviruses and potyviruses from the same NitroPure nitrocellulose membrane. Journal of Virological Methods, 171(2), 345–351.

Chen, H., Zhang, L., Yu, K., and Wang, A. (2015). Pathogenesis of Soybean mosaic virus in soybean carrying Rs v one gene is associated with miRNA and siRNA pathways, and breakdown of AGO1 homeostasis. Virology, 476, 395–404.

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, 183-194.

Chiquito-Almanza, E., Acosta-Gallegos, J. A., García-Álvarez, N. C., Garrido-Ramírez, E. R., Montero-Tavera, V., Guevara-Olvera, L., and Anaya-López, J. L. (2017). Simultaneous Detection of Both RNA and DNA Viruses Infecting Dry Bean and Occurrence of Mixed Infections by BGYMV, BCMV and BCMNV in the Central-West Region of Mexico. Viruses, 9(4), 63.

Cobos, A., Montes, N., López-Herranz, M., Gil-Valle, M., and Pagán, I. (2019). Within-Host Multiplication and Speed of Colonization as Infection Traits Associated with Plant Virus Vertical Transmission. Journal of virology, 93(23), e01078-19.

Eick, G., Urlacher, S. S., McDade, T. W., Kowal, P., and Snodgrass, J. J. (2016) Validation of an Optimized ELISA for Quantitative Assessment of Epstein-Barr Virus Antibodies from Dried Blood Spots. Biodemography and Social Biology, 62(2), 222–233.

Golnaraghi, A. R., Shahraeen, N., Pourrahim, R., Farzadfar, Sh., and Ghasemi, A. (2004) Occurrence and relative incidence of viruses infecting soybeans in Iran. Plant Disease, 88, 1069-1074.

Hartman, G.L., West, E.D., and Herman, T.K. (2011). Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Security, 3, 5- 17.

He, B., Hill, J. H. and Hajimorad, M. R. (2011). Factors to Improve Detection of Alfalfa mosaic virus in Soybean, Plant Health Progress, 12(1), 8.

Hill, J. H., and Whitham, S. A. (2014). Control of virus diseases in soybeans. Advances in virus research, 90, 355–390.

Jeong, J.J., Ju, H.J. and Noh, J. (2014). “A Review of Detection Methods for the Plant Viruses,” Research in Plant Disease. Korean Society of Plant Pathology, 20(3), 173-181.

Kimaru, S. L., Kilalo, D. C., Muiru, W. M., Kimenju, J. W., and Thuku, C. R. (2020). Molecular Detection of Cucumber Mosaic Virus and Tobacco Mosaic Virus Infecting African Nightshades (*Solanum scabrum* Miller). International Journal of Agronomy, 2020, 1-7.

Lamprecht, R. L., Kasdorf, G. G. F., Stiller, M., Staples, S. M., Nel, L. H., and Pietersen, G. (2010) Soybean blotchy mosaic virus, a new Cytorhabdovirus found in South Africa. Plant Disease, 94(11), 1348–1354.

Li, S. C., and Chang, Y. C. (2002). First report of Hibiscus chlorotic ringspot virus in Taiwan. Plant Pathology, 51(6), 803.

Li, Y. Y., Zhang, R. N., Xiang, H. Y., Abouelnasr, H., Li, D. W., Yu, J. L., and Han, C. G. (2013). Discovery and Characterization of a Novel Carlavirus Infecting Potatoes in China. PLoS ONE, 8(6), 1–13.

Marwal, A., Kumar Sahu, A., and Gaur, R. K. (2013). Molecular Characterization of Begomoviruses and DNA Satellites Associated with a New Host Spanish Flag (*Lantana camara*) in India. ISRN Virology, 2013(Ac, 1048), 1–5.

Nam, M., Bae, H., Hammond, J., Domier, L. L., Youn, Y.-N., Lee, B.-C., and Lim, H.S. (2013). Seed Transmission Rates of Bean pod mottle virus and Soybean mosaic virus in Soybean May Be Affected by Mixed Infection or Expression of the Kunitz Trypsin Inhibitor. Research in Plant Disease, 19(2), 114–117.

Niu, S., Gil-Salas, F. M., Tewary, S. K., Samales, A. K., Johnson, J., Swaminathan, K., and Wong, S. M. (2014). Hibiscus chlorotic ringspot virus coat protein is essential for cell-to-cell and long-distance movement but not for viral RNA replication. PLoS ONE, 9(11), 1–11.

Pietersen, G., and Garnett, H. M. (1990). Survey for the viruses of soybeans (*Glycine max*. L) in the Transvaal, South Africa. Phytophylactica, 22, 35-40.

Pietersen, G., Staples, S. M., Kasdorf, G. G. F., and Jooste, A. E. C. (1998). The relative abundance of soybean viruses in South Africa. African Plant Protection, 4(2), 65–70.

Pietersen, G., and Garnet, H. M. (2000). Soybeans in South Africa, 215(1992), 211–215.

Pietersen, G., and Morris, J. (2002). Natural Occurrence of Groundnut ringspot virus on Soybean in South Africa. Plant Disease, 86(11), 1271–1271.

Pflieger, S., Blanchet, S., Meziadi, C., Richard, M. M., Thareau, V., Mary, F., Mazoyer, C., and Geffroy, V. (2014). The "one-step" Bean pod mottle virus (BPMV)-derived vector is a functional genomics tool for efficient overexpression of heterologous protein, virus-induced gene silencing and genetic mapping of BPMV R-gene in common bean (*Phaseolus vulgaris* L.). BMC plant biology, 14, 232.

Rahbari, R., Moradi, N., and Abdi, M. (2021). rRT-PCR for SARS-CoV-2: Analytical considerations. Clinica Chimica Acta, 516, 1–7.

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.

Ramos-González, P. L., da Costa-Rodrigues, M., Potsclam-Barro, M., Chabi-Jesus, C., Banguela-Castillo, A., Harakava, R., and Freitas-Astúa, J. (2020). First genome sequence of

an isolate of hibiscus chlorotic ringspot virus from the Western hemisphere. Tropical Plant Pathology, 45, 153–158.

Rupe, J., and Luttrell, R. G. (2008). Effect of Pests and Diseases on Soybean Quality. Soybeans: Chemistry, Production, Processing, and Utilization, 2, 93–116.

Seo, J. K., Lee, Y. J., Kim, M. K., Lee, S. H., Kim, K. H., and Choi, H. S. (2014). A novel set of polyvalent primers that detect members of the genera Bromovirus and Cucumovirus. Journal of Virological Methods, 203, 112–115.

Shafie, R. M. (2019). Hibiscus chlorotic ringspot virus (HCRSV) affecting roselle (Hibiscus sabdariffa L.) in Egypt. Current Research Web. Middle East Journal of Applied Sciences, 9, 50–62.

Wang, R. Y. and Ghabrial, S. A. (2002). Effect of aphid behavior on efficiency of transmission of Soybean mosaic virus by the soybean-colonizing aphid, *Aphis glycines*, Plant Disease, 86(11), 1260–1264.

Zheng, L., Rodoni, B. C., Gibbs, M. J., and Gibbs, A. J. (2010). A novel pair of universal primers for the detection of potyviruses. Plant Pathology, 59(2), 211–220.

Chapter 3

Complete Genome Analysis of Hibiscus Chlorotic Ringspot Virus Isolate Infecting Soybeans in KwaZulu-Natal, South Africa

Abstract

Hibiscus chlorotic ringspot virus (HCRSV) is a positive sense single-stranded RNA virus that belongs to the family *Tombusviridae,* subfamily *Procedovirinae* in the genus *Betacarmovirus*. The virus was reported for the first time to infect soybean in South Africa (SA) in chapter 2. The aim of this study was to use Next Generation Sequencing (NGS) to analyse the complete genome sequence of HCRSV infecting soybean in the province of KwaZulu-Natal, SA. Total RNA extracted from soybean samples exhibiting virus-like symptoms was combined into one sample and used as template for NGS analysis. The sequence data generated was analysed using Genome Detective Virus Tool version 1.133. The HCRSV complete genome sequence obtained was compared with other HCRSV sequences from GenBank database using BLASTN. Pairwise and Multiple sequence alignments of the sequences were done using ClustalW tool available in MEGA X. Phylogenetic analysis was done using nine closely related HCRSV sequences including turnip crinkle virus (TCV) which was used as an outgroup. The open reading frames (ORFs) for the HCRSV genome were determined using ORF finder and protein sizes were measured using Protein Molecular Weight software. Recombination events were analysed using RDP4 software. NGS data analysis revealed that HCRSV, CMV and TMV were present in the infected soybean samples. Phylogenetic analysis results showed that the NdlovuNS_HCRSV-SA isolate from this study (Accession number: OK636421) was closely related to isolate XM from China with a bootstrap value of 99%. Genome organisation analysis of the NdlovuNS_HCRSV_SA isolate compared with other HCRSV isolates suggested high levels of similarity. The BLAST analysis correlated with the results from the genome organisation data, with the HCRSV isolates sharing 87.87% - 97.10% nucleotide identity. Recombination analyses showed a single event confirming that the NdlovuNS_HCRSV-SA isolate is a recombinant strain. To our knowledge, this is the first report of a full genome sequence of an HCRSV isolate infecting soybean in South Africa. Findings from this study are a first step towards understanding HCRSV on soybean in SA and developing effective strategies to manage the diseases it causes.

3.1. Introduction

Hibiscus chlorotic ringspot virus (HCRSV) is commonly found infecting hibiscus (*Hibiscus rosa-sinensis* L.*, H. cannabinus* L.*, H. sabdariffa* L.*, and H. trionum* L.) and ornamental plants worldwide (Huang et al., 2000; Zheng et al., 2018). The first discovery of HCRSV was in 1976, where the virus was found to naturally infect hibiscus plants in USA (Gao et al., 2012). HCRSV has a wide host range which includes some of the economically important plants which include tomato (*Lycopersicon esculentum* Mill.), cotton (*Gossypium barbadense* L.), tobacco (*Nicotiana glutinosa* L.), cowpea (*Vigna unguiculata* L.), common bean (*Phaseolus vulgaris* L.), pepper (*Capsicum annum* L.), eggplant (*Solanum melongena* L.), Rosella (*Hibiscus sabdariffa* L.), okra (*Abelmoschus esculentus* L.), and quinoa (*Chenopodium quinoa* Willd.) (Shafie, 2019). Some of the symptoms caused by HCRSV include vein banding, stunted growth, chlorotic ring spots, mosaic, dark red spot, necrotic local lesions, vein banding and mild chlorotic ring spot (Luria et al., 2013; Shafie, 2019; Li and Chang 2002).

According to the International Committee on Taxonomy of Viruses (ICTV) (2020), HCRSV belongs to the *Tombusviridae* family, subfamily *Procedovirinae* in the genus *Betacarmovirus.* HCRSV is transmitted mechanically or by vegetative propagation of infected plant, however it is not transmitted via seeds or aphid (*Myzus persicae*) (Shafie, 2019; Brunt and Spence, 2000). It has a positive sense single stranded RNA genome ranging from size 3.7 to 4.8 kb and encodes for seven proteins (ICTV, 2020). The ORF near the 5′ end encodes for p28 and its read-through p81 protein, both are involved in virus replication (Huang et al., 2000). This is followed by ORFs which are at the centre of the genome that code for two overlapping small proteins; the p8 and p9 movement proteins (Huang et al., 2000). The coat protein encoded by ORF4 is located near the 3′ end of the genome (Zhou et al., 2006). In addition, on the 5’ and 3’ ends of the genome, two ORFs encode for two unique proteins, p23 and p25 (Huang et al., 2000). More studies are still required to fully understand the functions of the proteins encoded by HCRSV.

Next Generation Sequencing (NGS) is now a commonly used tool in virus identification. NGS has the advantage that one does not need to have prior knowledge of the virus to be identified and it makes it possible to identify many viruses simultaneously. In NGS analysis, millions of sequenced data are generated by fragmenting the extracted DNA or RNA into numerous copies in a parallel manner. The procedure involves adding specialized adapters, sequencing the libraries, and reassembling them to form genomic sequences (Liu et al., 2016). A complete virus genome sequence can be obtained from the NGS analysis. This technique accurately identifies the entire virus population present in the sample and has ability to detect viruses even at low titres (Henson et al., 2012).

This study was a continuation of the study conducted in Chapter 2 of this Dissertation The results from the Chapter 2 study were thefirst report of HCRSV infecting soybean in SA. Against this background, the aim of this study was to perform a complete genome analysis of HCRSV infecting soybean in SA.

3.2. Materials and Methods

3.2.1. Sample Collection

As previously outlined in Chapter 2.

3.2.2. RNA extraction and Sequencing

Total RNA was extracted from 54 symptomatic soybean leaves using the Quick RNA™ MiniPrep Kit (Zymo Research, USA) and Quick RNA™ MicroPrep Kit (Zymo Research, USA) according to the manufacturer's instructions provided for each kit. A volume of 60 µl of the total RNA was eluted, quality checked, and the concentration was measured using a Nanodrop 1000™ spectrophotometer (ThermoFisher Scientific™, USA). The total RNA from all the 54 samples was combined into one sample for NGS analysis as a cut-costing measure. The combined sample was named AG3. The total RNA of sample AG3 was used as a template to carry out NGS analysis done at Agricultural Research Council - Biotechnology Platform (ARC-BTP), Pretoria, SA. 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 0.11.5, which is standard for NGS sequencing. The pair end of the data was analysed using the genotyping online pipeline software; Genome detective virus tool version 1.133 (Vilsker et al., 2019) with five methods embedded including 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 HCRSV genomes were selected and aligned using ClustalW found in Molecular Evolutionary Genetics Analysis (MEGA X, Kumar, et al., 2018) software to generate consensus sequences for the RdRp, p28, hypothetical protein, movement proteins P8 and P9, and the coat protein. The ORFs for each sequence were determined with the OFR finder in NCBI. The molecular weight

of the proteins was determined using the Protein Molecular Weight tool (https://www.bioinformatics.org/sms/prot_mw.html) (Stothard, 2000).

3.2.4. Phylogenetic analysis

The consensus sequences were subsequently subjected to the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/), which is commonly used in searching and aligning sequences. The sequence of the isolate NdlovuNS_HCRSV-SA (generated in this study) and HCRSV isolates from different countries obtained from the NCBI database were selected and aligned by Pairwise and Multiple alignmentsor the phylogenetic analysis (Table 3.1). The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model (Tamura and Nei, 1993). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree 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.

Table 3.1: Complete genome sequences of HCRSV isolates used in the phylogenetic analysis

Virus	Isolate Identity	Accession number	Host/cultivar	Country	References
HCRSV	NdlovuNS_HCRSV-SA	OK636421	Soybean (*Glycine max*)	South Africa	In this study
	XM	KY933060.1	*Hibiscus rosa-sinensis*	China	Zheng et al., 2018
	SBO1	MK279671.1	*Hibiscus rosa-sinensis*	Brazil	Unpublished
	OUGC	MT512573.1	*Hibiscus rosa-sinensis*	USA	Olmedo-Velarde et al., 2021
	–	X86448.2	*Hibiscus sp*	Singapore	Huang et al., 2000
	HCRSV-Is	KC876666.1	*Hibiscus rosa-sinensis*	Israel	Luria et al., 2013
	HCRV-UKM	MN080500.1	Hibiscus	Malaysia	Unpublished
	TW	DQ392986.1	*Hibiscus rosa-sinensis*	Taiwan	Li and Chang 2002
TCV [a]	–	NC_003821.3	–	USA	Carrington et al., 1989

Note: [a.] Turnip crinkle virus was used as an outgroup

3.2.5. Recombination Analysis

The recombination analysis was used to detect any recombination events which might have occurred between HCRSV isolates. The analysis was done using the recombination detection program (RDP) version 4.101 embedded with different statistical methods including 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). The nucleotide sequences for HCRSV isolates that were used for phylogenetic analysis were also used for the recombinant analysis (Table 2).

3.3. Results

3.3.1. RNA quality and quantity assessment, and Sequencing

The RNA concentration for sample AG3 was 429.0ng/µl with an absorbance A260/A280 ratio of 2.02. This ratioindicated the RNA was of high quality for use in NGS analysis. The total NGS raw data generated was 2.57 gigabytes according to the Genome Detection Virus Tool version 1.133. Before any further analyses, low-quality reads were trimmed. It started with 29664378 reads, and 1978892 reads (6%) that did not pass QC, were removed (Table 3.2). The pre-processing step filtered low quality reads and removed potential adapters and this allows for differentiating reads into viral and non-viral reads. Thereafter, the analysis started with 27685486 reads, then 27350576 reads (98%) that did not appear to be viral, were removed. The next step was de novo assembly from 334910 reads, with 9% of reads assembled into viral contigs and 31167 reads mapped back to these viral contigs. Then lastly, the identification of appropriate viral reference genomes to align contigs was done and HCRSV, CMV and TMV were identified in the combined sample AG3. The complete genome of HCRSV was retrieved and it had a coverage of 97%, and 1800 reads.

Table 3.2: NGS statistics

Characteristics	**Statistics**	
Size input files	1.21 GB	1.36 GB
Original length	151 bp	
Number of raw reads started with	29664378	
Reads that did not pass QC [a]	1978892	
Number of reads started before trimming	27685486	
Number of reads after trimming	27350576	
Average length	50 – 132 bp	
Reads started with de novo assembly	334910	
Reads mapped back to viral contigs	31167	

Note: [a] QC represents quality control

3.3.2. Phylogenetic Analysis

The complete genome sequence of Isolate NdlovuNS_HCRSV-SA (Accession number: OK636421) formed a cluster with the isolate XM (KY933060.1) from China with a bootstrap value of 99% (Fig. 3.1). This cluster was closely related to a cluster made up of isolate MK279671.1 from Brazil and isolate KC876666.1 from Israel with a bootstrap value of 66%. The phylogenetic tree was rooted using turnip crinkle virus (TCV) isolate (NC_003821.3).

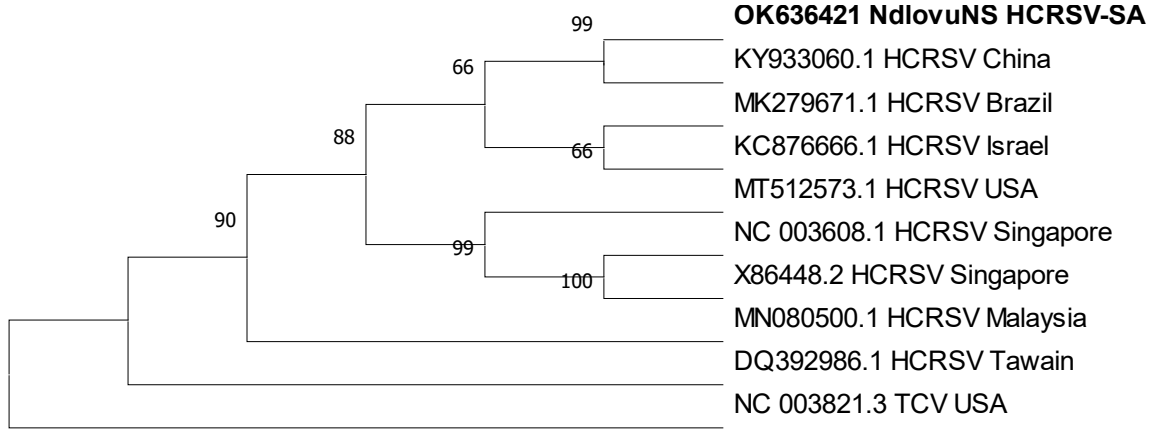

Figure 3.1: Phylogenetic analysis of HCRSV Isolates with complete genome sequences. The evolutionary analysis was inferred using Maximum likelihood (ML) and Tamura-Nei model. The bootstrap consensus tree was inferred from 1000 replicates.

3.3.3. Genome Organisation of HCRSV

The complete genome isolate NdlovuNS_HCRSV-SA was determined. The genome size was 3792 nt. It consisted of seven annotated ORFs; 5' proximal ORF (p28_1), located from 31 – 675 nt on the genome, encodes for P28 protein with a calculated molecular weight of 23.55 kDa. In the same ORF (p28_1), there is ORF (HCRSVgp1) located from 736 - 2238 nt on the genome. Using in-frame readthrough of the stop codon of ORF(p28_1), the reading frame extends towards the 3′ terminus of the genome, giving rise to the RNA-dependent RNA polymerase (RdRp) protein with a molecular weight of 57.14 kDa (Table 3.3). The ORF (p28_2) was identified located from 41 – 670 nt near the 5' proximal end of the genome and the ORF encodes for a hypothetical protein with a molecular weight of 23.72 kDa. Two overlapping ORFs (HCRSVgp3) located from 2205 – 2417 nt and ORF (HCRSVgp4) from 2257 – 2580 nt on the genome, encodes for movement protein P8 and P9 respectively. These two ORFs are located in the centre of the genome (Fig. 3.2). The 3′-proximal ORF (HCRSVgp5) from 2590 – 3627 nt on the genome, encodes for a protein with a molecular weight of 36.90 kDa. ORF (HCRSVgp6) located from 2570 – 3277 nt on the genome, encoded a hypothetical protein with 27.17 kDa molecular weight. This genome organisation of the NdlovuNS_HCRSV isolate is similar to that of other HCRSV isolates found on NCBI GenBank.

Table 3.3: Identification of 7 coding regions in the sequence alignment of NdlovuNS_HCRSV-SA (Accession number: OK636421) genome.

Gene	Protein	Genome region (nt) [b]	Length (nt) [b]	Protein (kDa)	aa [a]
HCRSVgp1	RdRP	736-2238	1503	57.14	500
p28_1	P28	31-675	645	23.55	214
p28_2	Hypothetical Protein	41-670	630	23.72	209
HCRSVgp3	P8	2205-2417	213	7.51	70
HCRSVgp4	P9	2257-2580	324	12.58	107
HCRSVgp5	CP	2590-3627	1038	36.90	345
HCRSVgp6	Hypothetical Protein	2570-3277	708	27.17	235

Note: [a.] aa- indicates amino acid

[b.] nt- indicates nucleotides

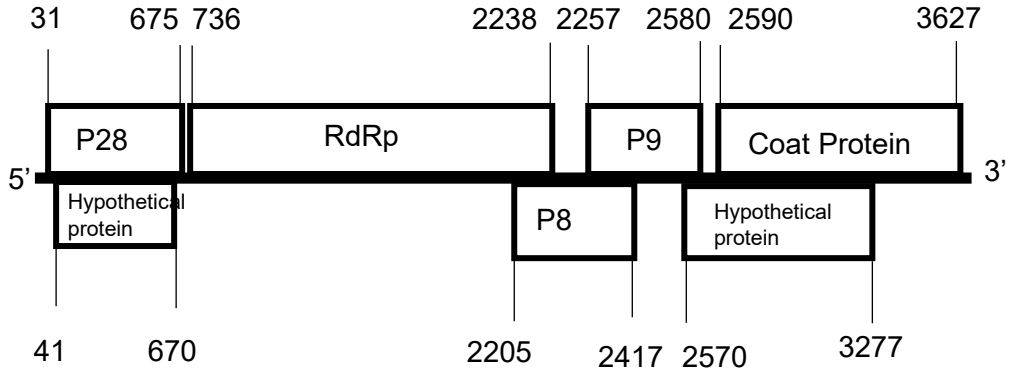

Figure 3.2: A schematic representation of the genome organisation of HCRSV, depicting coding proteins. The numbers included above and below the genome indicate the start position (nt) of each region.

The genome organisation data correlated with the nucleotides and deduced amino acid identities obtained by BLAST analysis. The analysis of the complete genomes showed that the HCRSV isolates shared 87.87% to 97.10% nucleotide identity (Table 4). The sequence of the XM isolate had the highest and isolate TW had the lowest nucleotide identity to NdlovuNS_HCRSV isolate. The amino acid identity data showed that isolate XM had the highest amino acid identity to the NdlovuNS_HCRSV isolate in six of the seven proteins compared to the other HCRSV isolates (Table 3.4).

Table 3.4: Comparison of nucleotide (nt) and deduced amino acid (aa) identities between HCRSV isolates

NdlovuNS_HCRSV-SA (OK636421)	HCRSV isolates XM (KY933060.1)	SBO1 (MK279671.1)	OUGC (MT512573.1)	HCRSV-Is (KC876666.1)	- X86448.1	HCRV-UKM (MN080500.1)	TW (DQ392986.1)
Nucleotide sequence							
Complete genome (nt)	**97.10%** [a]	94.23%	93.12%	93.12%	93.07%	92.73%	87.87%
Amino acid (aa)							
RdRp	**95.37%** [a]	93.06%	92.65%	92.79%	92.93%	91.69%	87.55%
P28	**91.12%** [a]	86.45%	85.51%	85.05%	87.38%	86.45%	78.97%
Hypothetical Protein	**77.22%** [a]	69.61%	70.56%	69.01%	68.89%	69.61%	58.56%
P8	**100%** [a]	97.14%	91.43%	97.14%	90.00%	97.14%	88.57%
P9	**98.15%** [a]	**98.15%** [a]	92.59%	94.44%	96.30%	96.39%	87.04%
Coat protein	**95.07%** [a]	94.78%	94.49%	91.59%	93.62%	93.91%	92.17%
Hypothetical Protein	88.84%	**91.07%** [a]	88.39%	86.61%	90.18%	88.84%	79.46%

Note: [a.] The highest values for the nucleotide identity and each amino acid identity are indicated with a bold face.

3.3.4. Recombination Analysis

A single recombination event occurred in isolate NdlovuNS_HCRSV-SA (OK636421) and was detected by RDP software version 4.101 (Martin et al., 2015). The event occurred from the nucleotide positions 2896 – 3089 in the alignment (Fig. 3.3). There were 5 methods embedded in the RDP program which confirmed this recombination event. These methods included GENECONV (Padidam et al., 1999) with 3.242 x 10-03, MaxChi (Smith, 1992) with 1.338 x 10-02, SiScan (Gibbs et al., 2000) with 1.665 x 10-01, BootScan (Martin et al., 2005) with 1.173 x 10-03 and 3Seq (Boni et al., 2007) with 5.103 x 10-02. The statistical significance with a P-value < 0.05 was accepted.

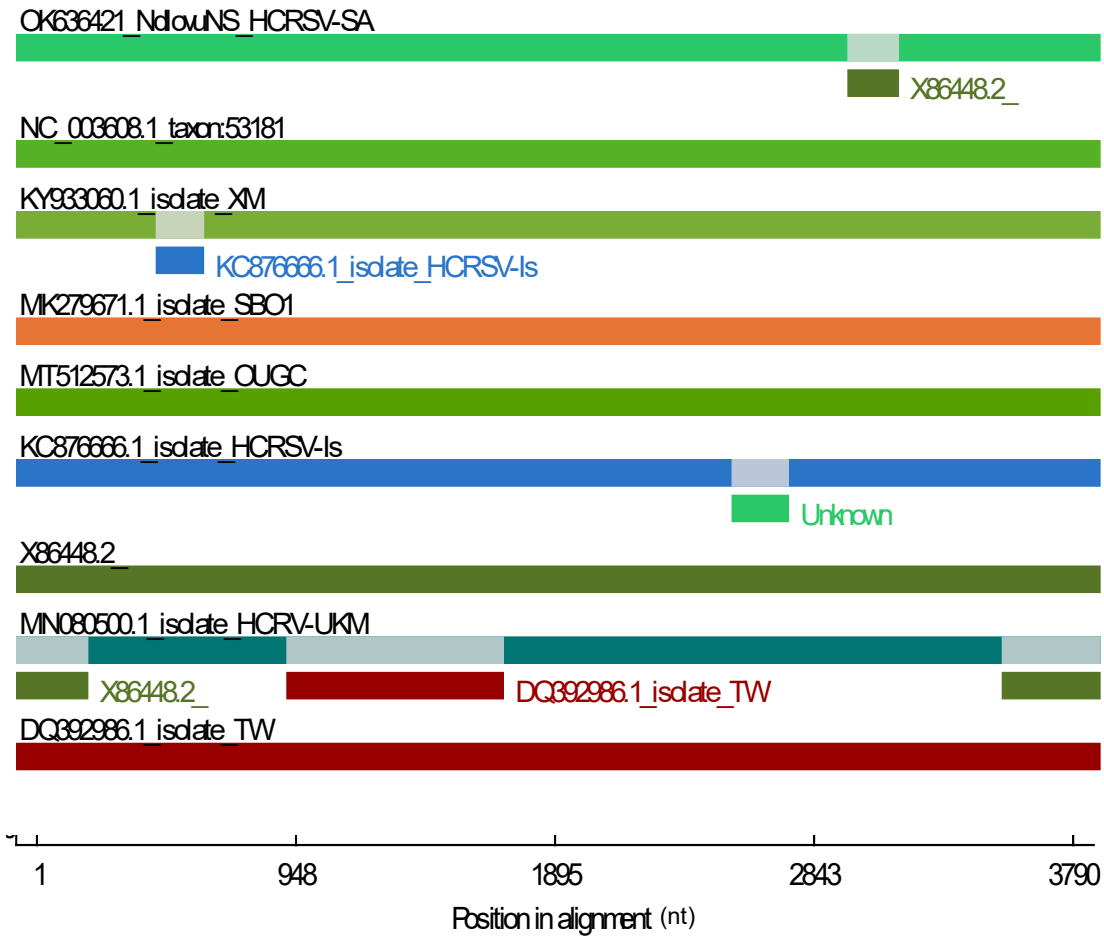

Figure 3.3: Recombination analysis of HCRSV isolates showing the regions where recombination events occurred. In the analysis, each isolate aligned are represented by a different colour. A number line was used to show the predicted recombination event in isolate NdlovuNS_HCRSV-SA which begins at nucleotide position 2896 and terminated at nucleotide 3089. The statistical significance was accepted, with a P-value < 0.05. The analysis was done using RDP 4 software version 4.101 (Martin et al., 2015).

3.4. Discussion

This study aimed to analyse the complete genome of HCRSV infecting soybean in KwaZulu-Natal province, South Africa. NGS generated high quantity data according to the NGS statistics and accurately detected three viruses in the soybean samples analysed. The complete genome sequence of HCRSV was obtained. Partial sequences for CMV and TMV were also obtained. This study only focused on HCRSV.

The sequence analysis of the NdlovuNS_HCRSV-SA isolate showed that it had a genome organization that was typical of HCRSV isolates from different parts of the world. Typically, HCRSV has seven ORFs (Huang et al., 2000; Rupe and Luttrell, 2008). All these

seven ORFs, together with the proteins they code for, were confirmed for the HCRSV isolate generated from this study (Table 3). The sizes of all the proteins of this isolate corresponded with those known to be coded for by typical HCRSV isolates (Figure 2). Further genome analysis showed high nucleotide and amino acid identities between the NdlovuNS_HCRSV isolate and some of the HCRSV isolates obtained from the GenBank database (Table 4). All this evidence points to the fact that the HCRSV generated from this study is not different from isolates of the virus whose sequences have been deposited in the GenBank database. To this end, our results confirm that this is the first report of HCRSV infecting soybean in SA. Previous studies have reported that the virus infects a number of plants but no reports were done on soybean plants (Shafie, 2019). We can conclude from these results that this is the first report of HCRSV infecting soybean worldwide.

Phylogenetic analysis showed that the isolate NdlovuNS_HCRSV-SA obtained from this study was closely related to the isolate XM from China with a bootstrap value of 99%. It is highly likely that infected plant material was exchanged between China and South Africa resulting in the virus recently appearing in South Africa for the first time.

RNA viruses rapidly evolve to form new strains of the virus (Rubio, 2013). Recombination plays a significant role in RNA virus genome evolution since it is responsible for the virus adaptation to infect new hosts (Kimaru et al., 2020; Rubio, 2013). In the present study, recombination analysis showed that the NdlovuNS_HCRSV isolate had one recombination event (Figure 3). The isolate was confirmed as a recombinant strain of the isolate from XM China (KY933060.1) with which it shared 97.7% nucleotide identity. It is highly likely that this recombination event was responsible for the adaptation of HCRSV to infect a new host (soybean) as reported in Chapter 2.

3.5. Conclusion

Our earlier study (Chapter 2) indicated that HCRSV was detected in the some of the soybean samples analysed. As a follow-up to that study, we analysed the complete genome of HCRSV isolate obtained in this study by NGS analysis. The results from the phylogenetic analysis showed that the NdlovuNS_HCRSV isolate is closely related to isolate XM from China. The genome organisation analysis also confirmed, in correlation with the BLAST analysis, results the NdlovuNS_HCRSV-SA isolate is closely related to isolate XM and other HCRSV isolates whose sequences have been deposited in the GenBank database. The recombination analysis results revealed that the NdlovuNS_HCRSV-SA isolate is a recombinant strain newly occurring in soybean plants which originated from the isolate XM from China and an unnamed Isolate from Singapore. According to our knowledge to date, this study reports for the first time the complete genome sequences of HCRSV isolate infecting

soybean South Africa. In addition, no reports of HCRSV infecting soybean has been reported from other parts of the world. Findings from this study are a first step towards understanding HCRSV on soybean in SA and developing effective strategies to manage the diseases it causes.

3.6. References

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., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology : a journal of computational molecular cell biology, 19(5), 455–477.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England), 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, 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 (Oxford, England), 33(19), 3137–3139.

Brunt, A. A. and Spence, N. J. (2000). The natural occurrence of Hibiscus chlorotic ringspot virus (Carmovirus; Tombusviridae) in aibika or bele (*Abelmoschus manihot*) in some south pacific island countries. Plant Pathology, 49 (6), 798-798.

Buchfink, B., Xie, C., and Huson, D.H., (2015). Fast and Sensitive Protein Alignment using DIAMOND, Nature Methods, 12, 59–60.

Carrington, J. C., Heaton, L. A., Zuidema, D., Hillman, B. I., and Morris, T. J. (1989). The genome structure of turnip crinkle virus. Virology, 170(1), 219–226.

Deforche, K. (2017). An alignment method for nucleic acid sequences against annotated genomes. bioRXiv. 10.1101/200394.

Gao, R., Ng, F. K. L., Liu, P., & Wong, S.-M. (2012). Hibiscus chlorotic ringspot virus Coat Protein Upregulates Sulfur Metabolism Genes for Enhanced Pathogen Defense. Molecular Plant-Microbe Interactions, 25(12), 1574–1583.

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, 573-582

Henson, J., Tischler, G., and Ning, Z. (2012). Next generation sequencing and large genome assemblies. Pharmacogenomics, 13(8), 901–915.

Huang, M., Koh, D. C.Y., Weng, L.J., Chang, M.L., Yap, Y.K., Zhang, L., and Wong, S.M. (2000). Complete Nucleotide Sequence and Genome Organization of Hibiscus Chlorotic Ringspot Virus, a New Member of the Genus *Carmovirus:* Evidence for the Presence and Expression of Two Novel Open Reading Frames. Journal of Virology, 74(7), 3149–3155.

Kimaru, S. L., Kilalo, D. C., Muiru, W. M., Kimenju, J. W., and Thuku, C. R. (2020). 'Molecular Detection of Cucumber Mosaic Virus and Tobacco Mosaic Virus Infecting African Nightshades (*Solanum scabrum* Miller). International Journal of Agronomy, 2020, 1-7.

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.

ICTV. (2020). International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/ accessed 11 October 2021.

Li, S. C., and Chang, Y. C. (2002). First report of Hibiscus chlorotic ringspot virus in Taiwan. Plant Pathology, 51(6), 803.

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), e63380.

Liu, J. Z., Fang, Y., and Pang, H. (2016). The current status of the Soybean-Soybean Mosaic Virus (SMV) pathosystem. Frontiers in Microbiology, 7(11), 1–9.

Luria, N., Reingold, V., Lachman, O., and Dombrovsky, A. (2013). Full-genome sequence of hibiscus chlorotic ringspot virus from Israel. Genome announcements, 1(6), e01050-13.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution,1, 1–5.

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 and Human Retroviruses 21, 98-102.

Olmedo-Velarde, A., Hu, J., and Melzer, M. J. (2021). A Virus Infecting Hibiscus rosa-sinensis Represents an Evolutionary Link Between Cileviruses and Higreviruses. Frontiers in Microbiology, 12, 660237.

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

Rubio, L. (2013). Genetic variability and evolutionary dynamics of viruses of the family Closteroviridae. Frontiers in Microbiology, 4, 151.

Rupe, J. and Luttrell, R. G. (2008). 'Effect of Pests and Diseases on Soybean Quality', Soybeans: Chemistry, Production, Processing, and Utilization, 2, 93-116.

Shafie, R. M. (2019). Hibiscus chlorotic ringspot virus (HCRSV) affecting roselle (Hibiscus sabdariffa L.) in Egypt. Current Research Web. Middle East Journal of Applied Sciences, 9, 50–62.

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

Stothard P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques, 28(6), 1102–1104.

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, 512-526.

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

Zheng, G. H., Liao, F. R., Ye, T., Zhang, W. Z., and Ming, Y. L. (2018). First Report of Hibiscus Chlorotic Ringspot Virus Infecting Hibiscus in Fujian Province, China. Plant Disease, PDIS–01–18–0117.

CHAPTER 4

General Overview

4.1. Major Findings

Plant viruses contribute to the economic losses of various crops on a global scale (Jones and Naidu, 2019). Knowledge about the viruses present in soybean (*Glycine max*. L) crops is important for implementing direct and effective control measures. Identification of 33viruses is the first step in devising appropriate measures for virus disease control. In South Africa (SA), few studies were undertaken many years ago on the identification of viruses infecting soybean. This was motivation to conduct this study, which was carried out with an aim to identify and characterise viruses infecting some soybean cultivars grown in KwaZulu-Natal, SA. In addition, determine the incidence of seed transmitted viruses. The information presented in this study from the experimental chapters is discrete; Chapter 2 and Chapter 3 will be sent for publication (after dissertation submission) in peer-reviewed journals, which will include first reports and a full research paper about the viruses infecting soybean in KwaZulu-Natal, SA.

The major findings generated in Chapter 2 of this study are based on detection and identification of viruses infecting soybean in KwaZulu-Natal, SA. CMV, TMV and HCRSV were reported for the first-time infecting soybean in SA. This chapter also included the first report of HCRSV infecting soybean worldwide. The results were confirmed by the Reverse Transcription Polymerase Chain Reaction (RT-PCR) which accurately detected and identified the viruses.

Chapter 3 was a continuation of the study done in Chapter 2, the study focused on the analysis of a complete genome sequence of hibiscus chlorotic ringspot virus obtained by the Next Generation Sequencing (NGS). High quality RNA was used as template for the NGS analysis and the total data generated was 2.57 gigabytes with 29 664 378 reads. The NGS results confirmed the identification of HCRSV, CMV and TMV reported in Chapter 2. The complete genome sequence of HCRSV obtained by NGS was analysed. The phylogenetic analysis showed that the isolate NdlovuNS_HCRSV-SA (OK636421) obtained from this study was closely related to the isolate XM from China (KY933060.1). The recombination analysis showed that the isolate is also a recombinant strain of the isolate XM from China.

4.2. Implications

The occurrence of HCRSV on soybean crops raises serious concerns for the soybean industry locally and globally. This virus may not be currently seen as a problem due to limited studies, however considering its wide host range which includes many economically important

crops, it can be a threat to soybean production. The symptoms for HCRSV include chlorotic ring spots, vein-banding, severe stunting and flower distortion on plants (Shafie, 2019; Raman and Muthukathan, 2015). According to this study, HCRSV causes stunted growth and vein clearing on soybean plants which can negatively affect the quality and quantity of soybean crops produced. As a result of yield reduction, the marketability of soybeans is also reduced and less sales are made out of it. However, conclusions cannot be drawn on the losses incurred by this virus on soybean due to limited studies. Furthermore, the identification of CMV and TMV on South African soybean is also concerning. These viruses have been known for a long time to cause major losses on various economically important crops (Kimaru et al., 2020; Salánki et al., 2018).

4.3. Way forward

The findings from this study showed that the virus structure in South African soybean has changed over time since the last study was conducted. More studies should be undertaken to further identify the viruses occurring in soybeans across SA. Surveys should be conducted in all provinces in order to identify the unknown/undetected viruses present in the soybean plants grown in SA. This can be achieved by using NGS; this technique has made it possible to identify many viruses simultaneously and obtain a complete virus genome sequence. Interestingly, this study reported HCRSV, CMV and TMV for the first-time infecting soybean in SA, and the results have opened a gap to further investigate these viruses on soybean. The study on complete genome sequence analysis for HCRSV needs to be done extensively to shed more light on how resistance can be developed against this virus. Soybean farmers must consider putting prevention measures in place to avoid crop losses incurred by the viruses reported in this study.

4.4. References

Jones, R. A. C., and Naidu, R. A. (2019). Global Dimensions of Plant Virus Diseases: Current Status and Future Perspectives. Annual Review of Virology, 6, 387–409.

Kimaru, S. L., Kilalo, D. C., Muiru, W. M., Kimenju, J. W., and Thuku, C. R. (2020). 'Molecular Detection of Cucumber Mosaic Virus and Tobacco Mosaic Virus Infecting African Nightshades (*Solanum scabrum* Miller)', International Journal of Agronomy, 2020, 1-7.

Raman, T., and Muthukathan, G. (2015). Field suppression of Fusarium wilt disease in banana by the combined application of native endophytic and rhizospheric bacterial isolates possessing multiple functions. Phytopathologia Mediterranea, 54(2), 241–252.

Shafie, R. M. (2019). Hibiscus chlorotic ringspot virus (HCRSV) affecting roselle (Hibiscus sabdariffa L.) in Egypt. Current Research Web. Middle East Journal of Applied Sciences, 9, 50–62.

Salánki, K., Gellért, Á., Nemes, K., Divéki, Z., and Balázs, E. (2018). Molecular Modeling for Better Understanding of Cucumovirus Pathology. Advances in virus research, 102, 59–88.

Conferences to date

Ndlovu, N.S and Gubba, A. (2019). Identification and Characterisation of Viruses Infecting Soybean (*Glycine max* L.) in South Africa. 3rd New Voices Symposium, in Pretoria, South Africa. 11 September 2019.

Appendix

Buffers for ELISA

- **PBS 1X [pH 7.4]**
  Dissolve in 1000 ml distilled water

NaCl	8 g
$Na_2HPO_4.H_2O$	2.9 g
KH_2PO_4	0.2 g
KCl	0.2 g
NaN_3	0.2 g

- **PBST 1X or washing Buffer**
  Add to 1000 ml PBS:

Tween20	0.5 ml

- **Extraction Buffer 1X [pH 7.3]**
  Add to 1000 ml PBS 1X:

PVP (Mw 10,000-40,000)	10 g
Tween20	0.5 ml

- **Coating Buffer 1X [pH 9.6]**
  Dissolve in 1000 ml distilled water

Na_2CO_3	1.59 g
$NaHCO_3$	2.93 g
NaN_3	0.2 g
Bromocresol purple	5 mg

- **Conjugate Buffer 1X [pH 7.4]**
  Add 1000 ml PBST 1X:

BSA	2 g
Congo Red	40 mg

- **Substrate (pNPP) Buffer 1X [pH 9.8]**
  Dissolve in 1000 ml distilled water:

Diethanolamine	97 ml
NaN_3	0.2 g