MOLECULAR TYPING OF WHEAT STREAK MOSAIC VIRUS FOR FORENSIC APPLICATIONS

By

JESSE DAVID CARVER

Bachelor of Science in Biological Sciences

University of Tulsa

Tulsa, OK

2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December 2007

MOLECULAR TYPING OF WHEAT

STREAK MOSAIC VIRUS FOR

FORENSIC APPLICATIONS

Thesis Approved:

Dr. Robert W. Allen

Thesis Adviser

Dr. Earl L. Blewett

Dr. Ulrich Melcher

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my Mom, Lois, and my Dad, Gene, for providing countless hours of care and support long before I considered pursuing a degree in higher education. What they have done throughout their lives has shaped me into the person I am today. My sister Jackie, my brother James, and my stepfather Dan have all played an important role in my life, and I surely would not be here without them.

I wish to express my heartfelt appreciation to my advisor, Dr. Allen, for all of his inspiration and dedication to a project and a graduate student who both, at times, appeared to be quite hopeless. I would also like to extend my appreciation to my committee members, Dr. Melcher and Dr. Blewett, who have both gone out of their way to provide me with a plethora of knowledge in a field where I initially was nothing more than a layman. Likewise, I would like to thank Dr. Jun Fu for all of her aid in the laboratory during the research time crunch. Without her assistance I would probably still be looking for certain reagents or protocols necessary for my project.

I would also like to show my appreciation for my friends and family, especially John Haley and Sara Duncan, that have helped me get my mind off-track when I needed it most so that I could continue to pursue this degree with a refreshed outlook and a renewed mind. Finally, without my fellow students Charlene Beauman, Dr. Gifty Benson, Kara Crosswy, Amber Hood, and Eric Duvall to make my time at the lab more enjoyable, my entire experience would have been rather dull and lifeless.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

Upon stepping down from his position as the United States Secretary of Health and Human Services, Tommy G. Thompson issued a grave warning about the future of U.S. agriculture: "For the life of me…I cannot understand why the terrorists have not attacked our food supply because it is so easy to do" (Pear, 2004). This advice is not without its merits as the economical impact of such an attack would be devastating. In 2008, the U.S. farming industry is projected to produce revenues of over $240 billion. Likewise, the United States Department of Agriculture, the main governing body of the US agricultural economy, is slated to have a budget of over $89 billion that will subsequently be reinvested into the enormous industry through programs and agricultural research (United States Department of Agriculture [USDA], 2007).

While these numbers seem staggering by themselves, fully accomplished revenues could only exist in a perfect world. Naturally occurring pathogens exist ubiquitously and significantly impact the world food market every year. Across the globe, "the damage from invasive plant, insect, and pathogen species is more than $100 billion annually" (Stack, 2006). Without aid or human intervention, roughly 3% of the world's crop and livestock is destroyed every year based on naturally occurring outbreaks and epidemics. It stands to reason then, that the effect of an intentionally released

pathogen, meant to directly disrupt the world-wide flow of farming products, could markedly increase the amount of unusable agricultural goods far above the unavoidable 3% suffered every year (USDA, 2007).

The successful intentional introduction of a plant or animal pathogen into a susceptible population could have countless economical repercussions for the victims. Despite the growing concerns however, such an attack would be somewhat difficult to accomplish. In theory, direct infection of a small portion of a field or small group of livestock could propagate an epidemic that would, given favorable conditions, wipe out an entire farm or ranch containing susceptible hosts. In practice, however, many uncontrollable elements exist:

> [I]t is difficult for man to initiate a pest or disease epidemic in nature. For a disease to develop, many interacting factors must occur in a coordinated manner; for an epidemic to result, those conditions must be sustained over time and space. The ability to regulate those factors to a degree that would promote an epidemic is lacking. (Stack, 2006)

Although many hurdles do exist, a savvy individual with the right knowledge, technology, and opportunity could intentionally introduce a limited outbreak.

Any outbreak has the potential to create far-reaching consequences. Monetary losses due to unsold crops would only be a short-term consequence; long-term losses under the shadow of bad press and potential agricultural quarantines could negatively impact the trade economy for years afterward. The majority of agroterrorism preparedness research strives to err on the side of caution and thereby become more defensive in nature. While a full epidemic might not be a possibility currently, as

technology and research continue to develop within the field of agriculture, so too will the potential threats associated with the development and spread of a given pathogen.

Wheat and wheat byproducts are the U.S.'s third largest agricultural export behind only feed grains and soybeans. Based on 2006 estimates, over $6 billion worth of U.S. wheat was sold outside of the country, accounting for nearly 10% of the total agricultural exports for the fiscal year (USDA, 2007). Wheat's global popularity stems from its many uses in daily life, from a staple food used to make flour to a component in fermentation of beer, vodka, and other alcohols. Based solely upon the world's dependence on the grain and its associated economic impact, wheat could conceivably be a crop targeted for a terroristic attack. An effective enough disruption of such a large market would have a ripple effect across the U.S.'s agricultural economy, serving the aims of the terrorist organization well.

While wheat is considered a naturally hardy crop, many pathogens use the grain as a host for their own development and replication (Watkins, 2002). One pathogen in particular, Wheat Streak Mosaic Virus (WSMV; family *Potyviridae*, genus *Tritimovirus*, species *aestivum* L) is a relatively ubiquitous virus that uses the wheat curl mite (*Aceria tosichella* [Keifer]) to gain horizontal passage to other wheat plants nearby (Stenger, 2005). Yield reductions due entirely to WSMV and the wheat curl mite are estimated to be at about 2% regionally, but localized losses of up to 100% can occur (Hunger, 1992; McNeil, 1996). Figure 1 shows a potyvirus, a family member of WSMV.

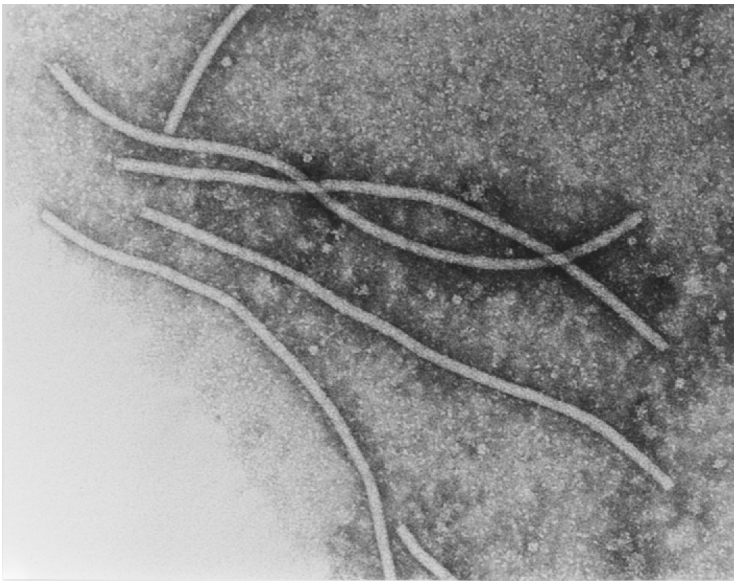


Figure 1. Electron micrograph of a potyvirus (Rothamsted, 1994)

The mite, WSMV's only known vector, is a wingless and tiny arthropod that spreads the virus by feeding on an infected plant; the virus takes up residence in the mite's saliva, and the mite can then transfer the virus to a new host upon its next feeding (Rabenstein, 2002). This viruliferous mite remains infected for roughly a week, and has the ability to continue to spread the disease to healthy hosts based upon its ability to feed on varying plant life nearby. Overall field infection based upon the mite's movements can be increased significantly with proper environmental factors like increased wind, high humidity, and moderate temperatures (Murray, 2005).

Because of its simplistic structure, WSMV lacks the mechanical ability to infect a plant without the aid of some wound-causing precursor. As the mite feeds though, the virus takes advantage of the abrasions created on the surface of the leaf (Burges, 2003). After this initial inoculation, the virus can migrate into the leaf and begin the process of infecting the new host. Figure 2 shows a scanning electron microscope image of the wheat curl mite.

Figure 2. Electron Microscopic Image of the Wheat Curl Mite (Townsend, 1996)

Once the mite has landed on the wheat leaf's surface, it begins to feed on the green plant material. While eating, the mite tends to stay on terminal leaves, and will migrate to each new leaf as it sprouts from the stem. Large populations of wheat curl mites feeding on a particular leaf will "cause the leaf margins to roll or curl inward hence the name" (Sloderbeck, 1995). Enough damage caused by the infestation will cause the wheat plant to dry down and the wheat mites that can no longer sustain themselves on the rapidly dying plant will begin to congregate on the flag leaves. As soon as the next gust of wind blows through the field, the mites will be transported off of the dying plant in hopes of attaching to a new, greener host. Thus, the cycle begins anew as the mite is now surrounded by more plant material, and it can restart the process of feeding (Sloderbeck, 1995).

Even without the virus entering the plant due to the mechanical abrasion brought on during feeding, the mite can cause a substantial amount of damage within a particular field. However, if the two work communally, the wheat plant stands little chance of

resisting the biological onslaught. While symptoms tend to vary widely based upon ambient temperature, general climate, wheat cultivar, and WSMV isolate, a particular set of characteristics are often associated with advanced WSMV infection (Montana, 1996). The infected plants are "normally stunted, and have mottled and streaked leaves. The leaf streaks are green-yellow and have parallel sides, and are discontinuous" (Townsend, 1996). Figure 3 shows a set of wheat leaves presenting full infection by WSMV.

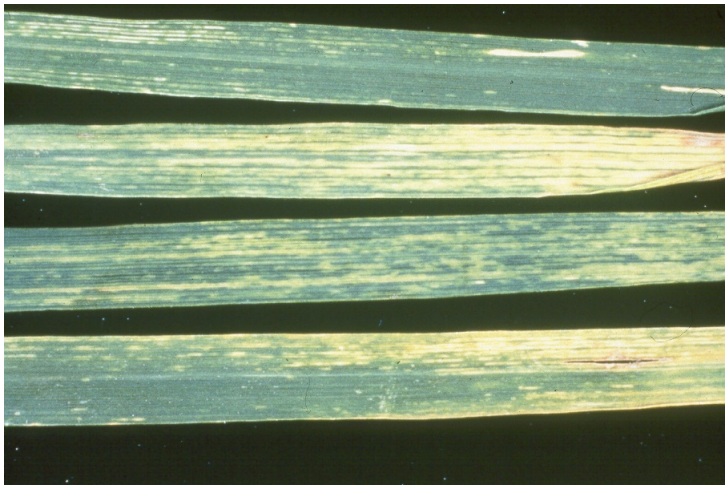


Figure 3. Wheat Leaf Infected with WSMV (Hammon, 2003)

In terms of evolutionary development, WSMV is fairly basic in its design. A single-stranded, positive sense RNA molecule is wrapped inside a non-enveloped capsid. In total, the entire virion measures less than 700 nm long (International Committee on Taxonomy of Viruses [ICTV], 2006), and the entire genetic code of the virus is only 9384 nucleotides long (GenBank, Accession #AF057533). This simplistic design, however, permits the virus to have a fairly wide host range encompassing many plants in the grass family; the virus can infect all varieties of wheat, and most isolates can infect barley, oats, maize, and millet (French & Stenger, 2003). WSMV's pathogenicity and virulence in contrast to wheat's susceptibility is fairly well researched (Chenault, 1996; French & Stenger, 2005; McNeil, 1996; Montana, 1996). In fact, before bioengineering

became popular and changed the outlook of the species as a whole, no wheat cultivar had been proven to have good resistance to WSMV (Sherwood, 1990).

The absence of WSMV resistant cultivars becomes vitally important when considering the propagation of the virus post-harvest. Being biotrophic, the virus cannot survive outside of a viable host or vector, and it cannot move or infect nearby hosts without aid from the wheat curl mite. The mite in turn requires leafy material for its continued existence, and even though the mite is resilient enough to survive for several days at below freezing temperatures, without food it too will die (Sloderbeck, 1995). With that understanding, both species' life cycles almost seem counterproductive to their respective abilities to live. Fortunately for the two species in question, a growing field does not remain dormant during the plant's off-season.

Volunteer wheat, a wheat plant that grows on its own rather than being deliberately planted, will thrive in the soil during the main wheat cultivar's non-growing season. These wheat plants actually serve as an oversummering host for the viruliferous mite, allowing both the mite and the virus to have a continual host even when the originally planted wheat cultivar is unavailable (Somsen & Sill, 1970). Despite the relative hardiness of both the virus and the mite, without the ability to exploit this constant source of susceptible plant life, neither would be able to survive the extended periods of time that occur between harvesting and planting. Without controls, the mite and virus population will continue to grow and will have a highly increased chance of infecting the new wheat very early in its development. This early infection proves to be very unfortunate for the growing wheat, as plants infected while young allow the virus

more time to develop and prove to be more severely damaged than those infected later in the growing period (Sill, Jr., 1953).

While time of inoculation during the growing period does play a major role in the development and onset of the disease, several aforementioned factors can also influence WSMV virulence: ambient temperature, general climate, wheat cultivar, and WSMV isolate (Montana, 1996). Temperature and climate's influence can easily be explained; pathogens prefer a specific environment and temperature range for the greatest activity. Many can survive and even replicate outside of their favored environment, but will do so at a level far from their abilities at peak virulence. Likewise, wheat cultivars can vary widely within their genomic structure to fit certain untapped niches within an environment. In other words, given enough time evolutionarily, a species can change so that every virus will not attack every cultivar with exactly the same intensity. In fact, cultivars are distinguished because of their internal variation, thereby further demonstrating the point.

Differences among WSMV isolates, however, are a bit harder to differentiate. The amount of genetic material per virus has already been mentioned as being minimal. In comparison, the human genome, recently sequenced by consortia, contains almost six billion nucleotides (Venter, 2001); WSMV's genome, on the other hand, contains less than ten thousand (GenBank, Accession #AF057533). This lack of genetic information does not allow much room for great variety within the species. However, like any creature placed under the pressure of evolutionary changes, WSMV would be forced to adapt on the strength of an often error-prone RNA replication process. Through successive replication, the RNA polymerase used to make copies of the viral genome will

inevitably make mistakes resulting in a rather high mutation rate for the species (Hall, 2001).

The most important evolutionary changes that a virus undergoes allow the virus easier access to the cell it wishes to infect. The virus's coat protein would be a suitable target for such mutations and can create differences in sequence sufficient to distinguish different isolates of the virus (McNeil, 1996). Through such changes, different isolates may evolve the ability to more effectively infect differing cultivars and plant species and therefore fill any open niches created by potential wheat-borne resistance. This assumption is backed by a number of recent research (via serology and Restriction Fragment Length Polymorphism mapping) that has found the virus's coat protein gene to harbor the most sequence heterogeneity within the genome (Montana, 1996; Chenault, 1996).

Slightly differing changes within the genes expressing the virus's coat protein could create shape differences on the surface of the capsid. Such external alterations could alter the ability of the virus to interact with a host cell membrane and, as a consequence, alter host range or the infection process (McNeil, 1996). However, due to the small size of the encapsulated virus any changes would require molecular analysis methods to discriminate among different isolates. If one were attempting to determine the modifications to genotype that characterize individual isolates, a deeper and more fundamental examination of the nucleotides at the heart of coat protein manufacturing is required.

Knowing that replication of the viral RNA genome can have errors suggests that a search of the genome could prove useful to distinguish strains. In this context, several

studies have shown that the majority of the WSMV genome is very stable (French & Stenger, 2005). The one region in which nucleotide sequence differences seem to accumulate is the gene encoding the viral coat protein, a region spanning 1213 nucleotides and located at the 3' end of the genome (Chenault, 1996). A coat protein gene of less than 1300 nucleotides does not allow for large genome changes without consequence for the very basic morphologies that define WSMV as its own species. The greatest variation, then, would inevitably have to occur in small, seemingly simple changes of the genome. A single nucleotide polymorphism (SNP) has the potential to diverge two closely related isolates without completely negating the functional characteristics of the viral capsid.

From a forensic viewpoint then, the question becomes how to develop or modify molecular methods that would be reliable and of sufficient discriminatory power to distinguish different strains of WSMV. Sequencing could be a viable option since the coat protein gene is less than 1300 nucleotides, a target that is manageable in 2-3 sequencing reactions. However, sequencing can be a cumbersome process, and the popular fragment analysis method employed by forensic laboratories operates with a different technology. Likewise, sequence output consists of nothing more than long strings of nucleotides that require in-depth study (or special alignment and comparison software) throughout to fully determine the comparative differences between isolates.

A more efficient and more reliable method would exploit a combination of simpler procedures that would each provide insight into the differences between any two given isolates. A process built on reverse transcriptase-polymerase chain reaction (RT-PCR) could fix the RNA into DNA, thereby reducing potential degradation that is often a

problem when working with RNA (McNeil, 1996). Then, an additional directly targeted process could allow researchers to visualize individual SNPs without having to perform an operation like sequencing which most modern forensic labs are ill-equipped to perform.

Of course, the entire process of identifying and targeting SNPs would mean nothing if the RNA reproduction mechanism was so faulty as to change the virus with each successive replication. It is known that "while a WSMV isolate may be defined by a consensus sequence, multiple and very closely related genotypic variants are [often] resident within a single isolate" (Stenger, 2002). Passage studies have confirmed that the known mutation rate (1-5nt/genome) does not consistently increase or decrease over the course of repeated rounds of replication (Stenger, 2002). Even though every consecutive virus is an additional generation further from the original founding population, mutation rates do not consistently spike and therefore the virus does maintain a general consensus sequence that is pervasive among all viruses of a particular isolate (French & Stenger, 2005). While the individual isolate genome is not particularly set, one can reasonably assume that certain changes are fixed within an isolate via consensus, while other changes are fleeting and may only exist on a per virus basis.

Obviously, the aforementioned fixed changes are, by far, the more important structural differences among the isolates. Knowing these changes to the consensus WSMV sequence (known as the Sidney-81 isolate) and being able to determine the differences for each virus encountered in the field would create possibilities of tracking the original viral generation. Finding a non-native isolate infecting a field would be

important, for sure, but proving that a suspected laboratory has cultured the same isolate of virus with a relative amount of scientific certainty would be substantially better.

In a world where the threat of bioterrorism and intentional infection of agricultural products through biological means is growing, having the ability to compare and effectively trace isolates of a particular disease would allow law enforcement to better perform their jobs. Samples collected from an intentionally infected field could then have direct forensic applications and could be compared to samples gathered from a clandestine laboratory suspected of cultivating the original virus. This practice would not exist without first having "the ability to distinguish WSMV isolates…[and therefore] enable direct comparison between those found in cultivated wheat and virus isolates prevalent in potential" clandestine laboratory samples (McNeil, 1996).

In order to evaluate the validity and potential forensic applications of molecular typing of WSMV isolates, the following research questions are posed:

1. Do distinct fixed SNP differences exist between different WSMV isolates?
2. Can WSMV be used to create an assay to distinguish these differences using current, popular forensic analysis equipment and protocols?
3. Will the assay efficiently and effectively discriminate between isolates based upon fixed SNPs while ignoring the more transitory, per virus changes?
4. Can an additional diagnostic step be attached to the assay's protocol as an internal control to guarantee that the virus under scrutiny is indeed WSMV?

CHAPTER II

MATERIALS AND METHODS

Wheat Streak Mosaic Virus Sample Collection and Storage

Three distinct WSMV isolates were used in this research. Two isolates, OSU and 964, had been previously collected or cultivated by Oklahoma State University; the third isolate, Plants A, B, C, and D, was gathered post-inoculation from living plants infected explicitly for the purposes of this research. Due to their common origin, Plants A-D are used interchangeably throughout the research.

Isolate 964 was residual virus collected in the field and maintained on frozen leafy material that had been placed in liquid and stored in a -20°C freezer for an indeterminate, but extended period of time. The OSU isolate was received as purified viral RNA in solution obtained from the plant pathology diagnostic lab at Oklahoma State University in Stillwater. Virus harvested from plants A through D constitute the third isolate; said virus was the result of an *in vivo* infection of four germinating wheat plants in a process that will be described in further detail in the following section.

Considering their varying compositions, all three viruses, when not in use, required different methods of storage based upon their media:

- Isolate 964 was left in its solution and stored in a -20°C freezer.

- Due to the increased potential for RNA degradation, isolate OSU was left in solution and stored in a -70°C freezer.
- Infected leaves from Plants A-D were cut and placed dry, in a -70°C freezer, as this is the generally preferred method of long-term storage (Gould, 1999).

Inoculation of Healthy Plants with WSMV

Two rounds of inoculation were completed in the course of this research. In each instance, three wheat plants were seeded and allowed to germinate roughly two weeks before attempting inoculation; two of the plants were designated for infection while the third would stand as a control and be allowed to grow without any contact with viral material. Plants A and B are products of the first round of inoculation; to ensure the proper virus was cultivated, the virions used to infect both plants were received as infected leaves from Dr. Roy French (University of Nebraska's Plant Pathology Department). Plants C and D are subsequent passages of infectious material gathered from Plants A and B after they had presented symptoms of WSMV infection.

To intentionally infect the newly sprouted wheat plants, several of the leaves showing evidence of infection with WSMV were combined with diatomaceous earth and water. The combined ingredients were rendered into a slurry via mortar and pestle. Once the solution was sufficiently blended, a droplet of the solution was rubbed lightly across the adaxial surface of each leaf in the growing plant.

Since WSMV is an opportunistic virus and therefore lacks the ability to infiltrate the leaf's defenses mechanically, the diatomaceous earth acted as an abrasive to create minute lacerations across the leaf. These injuries served as a point of penetration for the

virus so that it could begin infection with subsequent growth and reproduction. Once treated, the plants were placed in a greenhouse and given two months to allow the virus to fully infect the plant.

Viral RNA Extraction

Viral RNA extraction was achieved through the use of a kit available from Ambion, Inc. The kit, MagMAX™ AI/ND Viral RNA Isolation Kit, specifically targets RNA while eschewing DNA through the use of magnetic bead technology. Due to the large available binding surface and the ability to be fully dispersed in solution, the beads are meant to work exceptionally well to allow thorough RNA binding, washing, and eluting.

Cut sections of fully WSMV infected leaves were placed in a mortar with 1.5mL of water and crushed with a pestle to create a viscous solution of plant and viral products. The solution was then gathered via micro-pipette to retrieve solely the liquid portion; the solid portion of the plant clustered together and was removed by hand from the solution and discarded. 400µL of liquid sample were added to 802µL of Viral Lysis/Binding Solution, as per the kit's instructions. 20µL of Bead Resuspension Mix were added to this solution and mixed for four minutes to allow lysing of the viral cells while the beads are fully dispersed in solution to bind their complement of viral RNA.

The beads were captured via a magnetic stand, the supernatant was aspirated, and the beads were subjected to three separate wash phases comprised of two different solutions included with the kit (400 µL of Wash Solution 1 once, followed by 200 µL of Wash Solution 2 twice) for 30 seconds per step. Finally, the beads were immersed into

50µL of warmed (65℃) Elution Buffer (included with the kit) and allowed to sit at room temperature for three minutes. The beads were recaptured on the magnetic stand and the eluent was transferred into an RNase-free processing tube. Unless progressing directly to the next procedure, the newly captured and purified viral RNA was placed in a -70°C for storage.

cDNA Synthesis and Amplification

Upon completing the extraction and purification of the viral RNA, specific areas of interest needed to be converted to cDNA and then amplified to create enough copies to allow further experimentation. The labile nature of viral RNA requires stabilization for its viability in any long-term research. Creating stable viral genomic sequence was aided by the use of a kit available from Invitrogen, Co. called SuperScript™ One-Step RT-PCR with Platinum© *Taq* (error rate: 1.8×10^{-6} substitutions per nucleotide per cycle) that specifically converts RNA into cDNA via reverse transcriptase (RT) before using polymerase chain reaction (PCR) to amplify selected genes in the RNA transcript (Monie, 2005).

Primers, therefore, would be needed to establish exact points to begin and end amplification of the viral genome. Through previously completed research, three specific primers, RCF1 (5'-AGCTGGATCCTTTTTTTTTTTTTT-3'), XC1 (5'-AACCCACACATAGCTACCAAG-3') and XV1 (5'-GTTGAGGATTTGTACTT-3'), were synthesized to complete this task (McNeil, 1996; Chenault, 1996; French, 2005). RCF1 is the RT-PCR primer that works to synthesize complementary strands of the ssRNA viral genome to create cDNA; XV1 and XC1 are sense and anti-sense primers,

respectively, that bind to the cDNA and amplify a 1371 nucleotide region that contains the entire coat protein open reading frame and flanking regions (French, 2005).

The parameters of the RT-PCR were followed according to the protocol provided with the SuperScript™ kit. The PCR reaction mixture is detailed in Table 1.

Table 1. Reaction Components and Volumes for SuperScript™ RT-PCR Amplification

Component	Volume/50µL
2X Reaction Mix	25 µL
Template RNA	10 µL
Sense Primer, XV1, (10 µM)	1 µL
Anti-sense Primer, XC1, (10µM)	1 µL
RT-PCR Primer, RCF1, (10 µM)	1 µL
RT/Platinum© *Taq* Mix	1 µL
Autoclaved distilled water	11 µL

The PCR reaction tubes were then placed in a thermal cycler and allowed to amplify using a cycling program defined by both the Superscript™ kit protocol and also conditions predicted by the specific primers in use. Table 2 below details the specifics of the thermal cycler program used in conjunction with the SuperScript™ kit.

Table 2. Thermal Cycler Program of SuperScript™ RT-PCR

cDNA Synthesis		Amplification			Elongation
1 cycle		40 cycles			1 cycle
Synthesize	Deactivate	Denature	Anneal	Replicate	Elongate
60 min.	2 min.	30 sec.	30 sec.	1 min.	5 min.
50°C	94°C	94°C	55°C	68°C	68°C

Gel Electrophoresis

To ensure the success of the RT-PCR amplifications, each RT-PCR product was subjected to gel electrophoresis. A 20 well, 1.5% agarose gel was created and was electrophoresed in 100mL of 1X Tris-Acetate EDTA (TAE; 10mM Tris-acetate (pH 8.3) and 1mM EDTA) running buffer.

Individual gel lanes included running samples of both RT-PCR products and DNA size standards. The two DNA size standards used, *BstEII* fragments of lambda phage DNA and *HaeIII* fragments of PhiX-174, were chosen solely on their compatibility to the expected fragment size of the amplicons. Figure 4 shows the bands created in a gel by both DNA size standards.

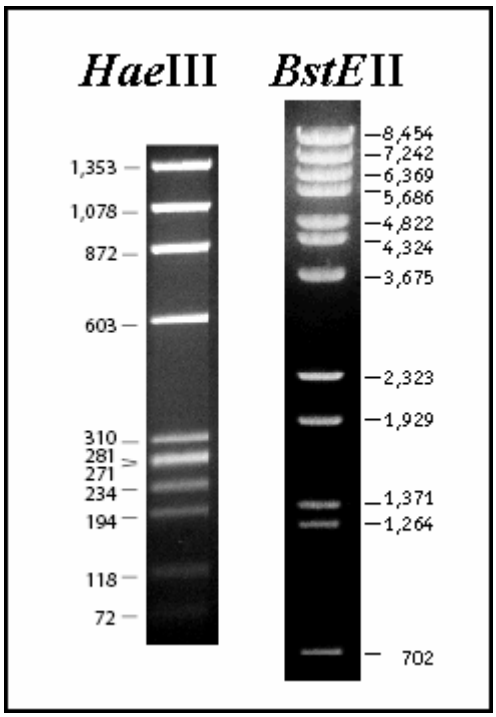


Figure 4. DNA Size Standards Used during Gel Electrophoresis

10µL of each 50µL RT-PCR sample were added to 10µL of 1X TAE loading dye containing bromphenol blue and xylene cyanol; likewise, the DNA size standards were mixed with 1X TAE loading dye at a ratio of 1µL to 19µL respectively. The samples were loaded onto the gel under 1X TAE running buffer and were electrophoresed for 105 minutes at 65V.

Once the samples had sufficient time to migrate, the gel was placed in a bath of 1X TAE running buffer and 200µL of a solution of ethidium bromide (EtBr) composed of 0.2g EB powder in 50mL of H_2O was added. The gel was incubated at room temperature in the TAE/EtBr bath for seven minutes to allow thorough staining to occur. The gel was then placed under lukewarm water for seven minutes to remove any unintercalated EtBr that might complicate visualization. Finally, the gel was placed on an ultraviolet light source to visualize the nucleic acid bands.

Zymo Column Purification

Removal of the unused nucleotides is vital for any additional experimentation involving any specialized form of PCR. Doing so requires the use of a specific column or filter targeted at removing smaller products. The Zymo Column/DNA Clean and Concentrator™ Kit was used to separate the sample into two distinct parts: purified cDNA and additional, unwanted PCR reagents.

The Zymo Column was first loaded with a mixture of two parts DNA Binding Buffer and one part DNA sample. This mixture was placed into a centrifuge and spun at a high speed (≥10,000 x g) causing the solution to pass through a filter column and into a collection tube. The desired DNA was now bound to the column while the effluent was discarded. The column was treated to a wash cycle twice; in both cases, 200µL of Wash Buffer was added to the column, and the column was centrifuged again at ≥10,000 x g.

The repeated rounds of washing remove contaminants and effectively clean and purify the cDNA. The column was then subjected to two rounds of elution by forcing 20µL of TE^{-4} (10mM Tris-CL, 8.0pH, 0.1mM EDTA) warmed to 70°C through the column at the same high rate of speed as before. Prior to the centrifugation step, however, the warmed TE^{-4} was allowed to sit within the column while the entire tube was placed on a 70°C heating block for one minute. While not absolutely necessary, this step gives the elution buffer a longer amount of time in the preferred environment to more thoroughly remove the DNA that had previously been bound to the column. The eluate was then transferred to a smaller collection tube and stored at 4°C until needed for other methods.

Amplified Fragment Length Polymorphism

In an attempt to distinguish differences between isolates, the amplicons were subjected to Amplified Fragment Length Polymorphism (AFLP). AFLP's technology takes advantage of the cutting sites associated with known restriction endonucleases and subjects long lengths of nucleic acid to these enzymes. The enzymes bind at very specific points in the DNA and sever the bonds that hold the DNA together at that point only. With the proper restriction endonucleases, the DNA will be cut into distinct lengths of nucleic acid which can then be separated from one another through electrophoresis.

Since the length of cDNA amplified by the RT-PCR process is relatively small, restriction enzymes with four-base and less frequently cutting six-base recognition sites would provide the best distinction between isolates. Figure 3 provides the recognition site of the two restriction enzymes used: *Hpa*II and *Ava*I.

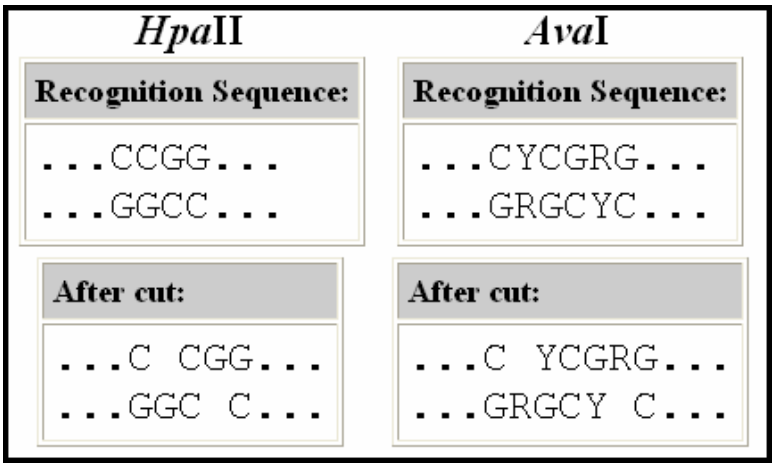


Figure 5. Recognition Sites of *Hpa*II, *Ava*I. An R in a recognition sequence stands for either adenine or guanine. A Y stands for thymine or cytosine. (Lab Rat, 2005)

For the reaction to work properly, both *Hpa*II and *Ava*I must be added to RT-PCR amplicons for a given period of time; however, both restriction endonucleases work optimally in different solutions and cannot be used simultaneously because of one main solution component: NaCl. While the two restriction endonucleases favor similar environments among other characteristics, the preferred salt concentration for *Hpa*II is much lower than the same for *Ava*I so, the cDNA was digested with *Hpa*II first.

10µL of cDNA amplicons, 2µL of RE_{ACT}© 2, *Hpa*II's assay buffer (available from Fermentas Life Sciences, Inc., Burlington, ON), 1µL of *Hpa*II enzyme (at a concentration of 5U/µL), and 7µL of autoclaved distilled water were combined in a processing tube and placed in a 37°C water bath for one hour. After the heating phase, 4µL of RE_{ACT}© 8, *AvaI*'s assay buffer (available from Fermentas Life Sciences, Inc., Burlington, ON), 1µL of *Ava*I enzyme (at a concentration of 10U/µL), and 15µL of autoclaved distilled water were added to the original processing tube's contents and the tube was returned to the 37°C water bath for an additional hour. Following the second heating phase, the newly cleaved DNA amplicons were now ready to be separated via gel electrophoresis.

The analyzed sample consisted of 20µL of the double digest combined with 5µL of 5x TAE loading dye. This sample was electrophoresed in a 1.5% agarose gel alongside the same DNA size standards as before (*HaeIII* and *BstEII*) for 105 minutes at 65V. Afterward, 200µL of EtBr were used in the staining and destaining process as previously described.

DNA Sequencing Methods

The most common method of DNA sequencing is the dideoxy chain-termination/Sanger method (Ziebolz & Droege, 2007). This process of sequencing is much like PCR-based DNA replication except for one distinct difference: instead of generating fully formed duplicates that are exact copies of the original template, truncated DNA strands are produced. This truncation occurs when a dideoxynucleotide (ddNTP) is inserted into a growing DNA backbone instead of a deoxynucleotide (dNTP).

Deoxynucleotides are the building blocks that form the backbone of DNA and are added normally to a growing DNA chain; they have a 3'-hydroxyl group on their deoxyribose sugar that allows further elongation of the DNA chain by supporting a phosphodiester bond with other dNTPs. Dideoxynucleotides, as their name would suggest, lack this important 3'-hydroxyl group, and, although they can be added to a growing DNA chain once at the end of the chain they will not allow any further nucleotides to be added. Frederick Sanger, a biochemist, understood the importance of ddNTPs and developed a method of DNA sequencing that capitalized on ddNTPs' inability to facilitate further DNA chain growth. His method suggested using a reaction mixture that heavily favored dNTP over ddNTP to the ratio of 99 to 1, respectively; this ratio allowed the DNA to begin replication, but 1% of the time the expanding chain would terminate due to the addition of a ddNTP.

Samples run using Sanger's methods contain an assortment of prematurely terminated DNA chains of differing sizes. These chains can be separated based on size through the use of electrophoresis and can subsequently be visualized with the aid of fluorescent labeling of one of the component parts used to build the replicated DNA.

Two types of experiments based on Sanger's originally developed principles were performed during this research; however, both experiments differed from one another based on the extent to which DNA sequencing was performed.

To gain a better understanding of the viral genomes, knowledge of their individual differences was crucial to aid in developing any discriminatory assays. Elucidating these differences required the sequencing of the coat protein region of all three isolates. This step was performed under contract by Eton Biosciences, Inc. (San Diego, CA). Plant A, Isolate 964, and OSU RNA was extracted with MagMAX and amplified with SuperScript producing a cDNA template; the template was sequenced with the coat protein region flanking primers (XC1 and Carver-1—5'-TACTTGACTGGGACCCGAA-3'; a sense primer analogous to XV1 that binds adjacent to it within the viral genome) as starting points. The results were compared by hand and through the software Spidey, an online mRNA to genomic alignment program available from NCBI, to determine possible SNPs.

Single Nucleotide Sequencing

One of the main goals of this research was to identify nucleotide sequence differences between WSMV isolates. As previously mentioned, the changes within the isolates occur on a very small, single nucleotide basis. One experimental approach envisioned developing a method that incorporates Sanger's original protocols but does so for only one of the four nucleotide bases. To determine the effectiveness of this approach, experiments were conducted that targeted only the guanine residues for early ddGTP termination, and did so by using fluorescently labeled sequencing primers.

In one approach, the Carver-1 primer that had been labeled with the fluorescent molecule FAM on the 5' end of the oligonucleotide, was used to initiate nucleotide sequencing. The PCR reaction mix consisted of the following: 10µL of 2mM Reaction Mix with Buffer (supplied by Invitrogen as part of their SuperScript™ Kit), 2.0µL of cDNA template, 0.5µL of *Taq* polymerase, 1.0µL of 1:100 dil of a mixture of 10mM unlabeled dGTP/ddGTP, 0.32µL of 10mM FAM-labeled Carver-1 primer, and 6.18µL of autoclaved distilled water.

The reaction contained a ratio of 5% ddGTP to dGTP in an attempt to force more truncated DNA chains to exist earlier in the reaction. While not working as optimally as 1% ddGTP to dGTP, this increased ratio successfully increased the rate of early truncation and would appear more readily once the DNA fragments were visualized and serve as a litmus test of future experimental constraints. Once the reaction was proven successful through capillary electrophoresis, the ratio would slowly be decreased to expand the potential length of the prematurely ended DNA chains so that more information could be gathered from the test.

The second type of single nucleotide sequencing experiment used a fluorescently labeled ddGTP instead of a fluorescent primer. 10µL of 2mM Reaction Mix with Buffer (again, supplied by Invitrogen as part of their SuperScript™ Kit), 2.0µL of cDNA template, 0.5µL of *Taq* polymerase, 1.0µL of 1:10dil 1mM Fluorescein-12-ddGTP, 0.32µL of 10mM Carver-1 primer, and 6.18µL of autoclaved distilled water composed the reaction mix for the second experiment. Again, a 5% ddGTP to dGTP ratio was employed initially to provide more definitive proof of the experiment's success. Both experiments were subjected to the PCR program described in Table 3.

Table 3. Thermal Cycler Program of ddGTP Sequencing

Initial Denaturing	Amplification		
1 cycle	25 cycles		
Denature	Denature	Anneal	Replicate
2 min.	30 sec.	5 sec.	4 min.
96°C	96°C	50°C	60°C

After completing the PCR cycle, the products were stored at 4°C until they were purified via the Centri-SEP method as described in a following section.

Complete DNA Sequencing

While single nucleotide sequencing targets a specific dNTP for early termination, complete DNA sequencing uses all four available ddNTPs to create truncated DNA chains that are terminated at every available nucleotide residue. Complete DNA sequencing was achieved through the use of a kit called the BigDye® Terminator v1.1 Cycle Sequencing Kit available from Applied Biosystems, Inc. (Foster City, CA). The kit uses a pre-mixed ratio of ddNTP to dNTP that allows the PCR-based extension of the DNA template to proceed while randomly attaching a ddNTP and terminating further chain elongation; the intent is to have ample amounts of sequencing truncation so that every base pair will fluoresce its appropriate color, creating a nucleotide by nucleotide sequence.

The reaction mixture consisted of the following: 4µL of Ready Reaction Premix, 2µL of BigDye Sequencing Buffer, 0.32µL of 10mM Carver-1 primer, 2µL of template DNA from the PCR amplicons, and 11.68µL of autoclaved distilled water. This mixture was then placed in the thermal cycler and run on the program described in Table 4.

Table 4. Thermal Cycler Program of BigDye® Terminator Sequencing

Initial Denaturing	Amplification		
1 cycle	25 cycles		
Denature	Denature	Anneal	Replicate
1 min.	10 sec.	5 sec.	4 min.
96°C	96°C	50°C	60°C

Once the BigDye cycle sequencing products had completed the thermal cycler program, they were stored in a 4°C refrigerator until they could be purified via the Centri-SEP method as described in a following section.

Centri-SEP Column Purification

A purification step is required following many PCR methods; this is especially true when fluorescently-linked molecules are included in the reaction mixture. Without a cleaning step, unused fluorescence will shine indiscriminately, making any additional visualization methods complicated by the overabundance of background luminescence. A badly formed baseline, unexpected abnormally high peaks, or unnaturally low expected peaks could all result without the proper clean-up of the PCR products. Kits specifically tailored for removing, or at least minimizing fluorescence from unincorporated fluorescence, are available and, if used, greatly simplify interpreting results.

After performing DNA sequencing, the amplicons were purified via separation through a CENTRI-SEP Column (available from Princeton Separations, Inc., Adelphia, NJ) that directly targets DNA fragments, separating said nucleic acids from labeled dNTPs/ddNTPs. 0.80mL of reagent grade water was added to a pre-mixed gel powder to

reconstitute the column. After allowing the gel two hours to hydrate, the column was centrifuged for 2 minutes at 750 x g to remove any remaining liquid from the resin. 20µL of fluorescent PCR amplicon was added directly to the center of the gel bed, and the column was placed in a collection tube and spun for another 2 minutes at 750 x g. The purified sample pooled at the bottom of the collection tube; the contents of said tube were placed at 4°C until analysis via the ABI 310 Genetic Analyzer could be performed.

SNaPshot Analysis Method

Given that SNPs are the main polymorphisms likely to distinguish WSMV isolates, a PCR-based, SNP analysis kit from Applied Biosystems, Inc. (Foster City, CA) called the PRISM® SNaPshot™ Multiplex Kit was investigated as a molecular typing method for WSMV. SNaPshot™ relies on primers that anneal to the template DNA one nucleotide upstream from a known SNP. Since the reaction mixture contains only ddNTPs, when the normal elongation phase of PCR begins, a single, fluorescently-linked ddNTP is added to the primer that complements the template DNA at the SNP site in the genome. Since the sizes of the primers are known, further visualization through the use of capillary electrophoresis in the ABI 310 Genetic Analyzer will reveal a colored peak one nucleotide larger than the size of the primer, and fluorescent color emitted will show which genotype exists at the SNP site. Figure 6 provides an overview of the SNaPshot™ process.

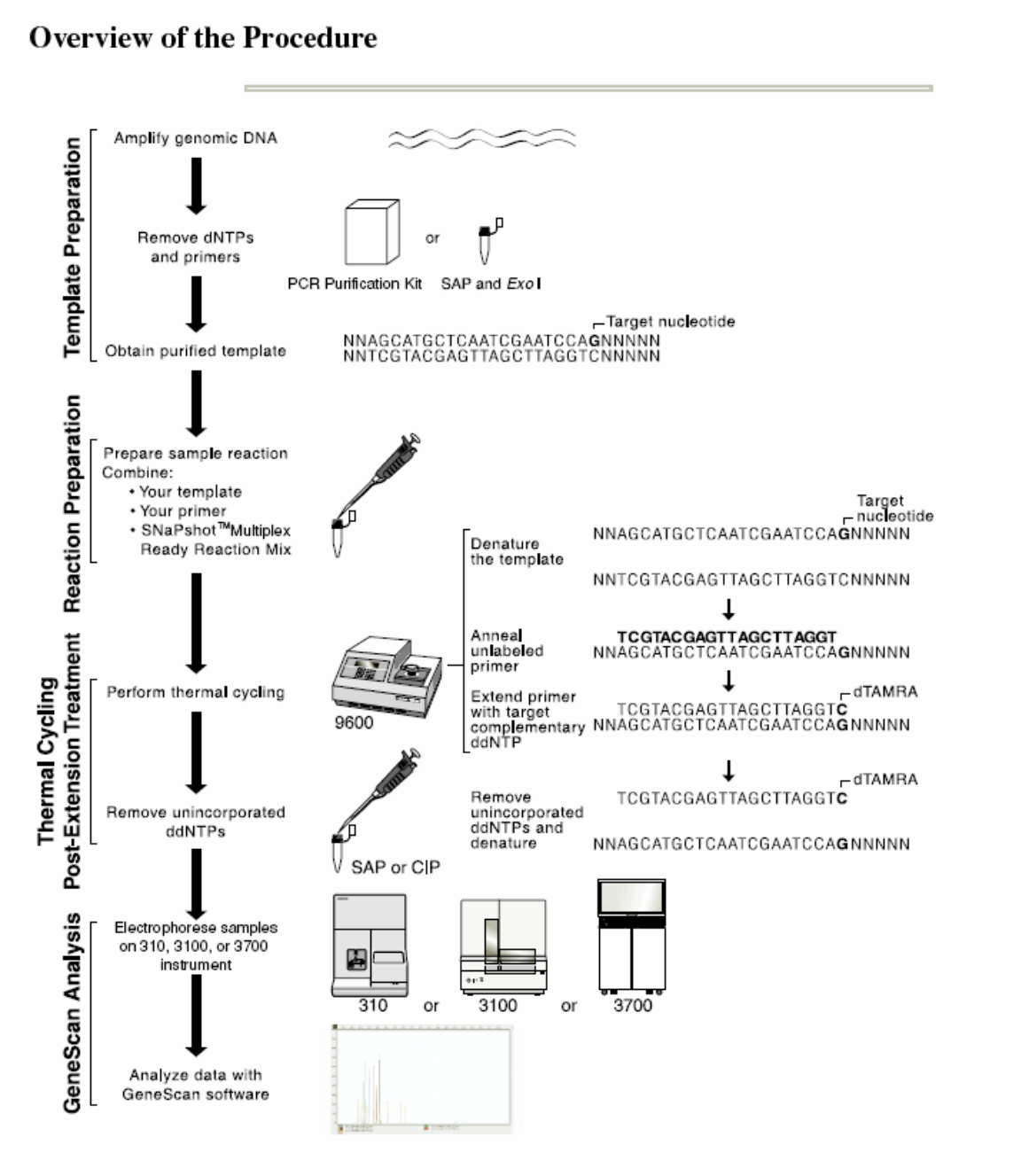
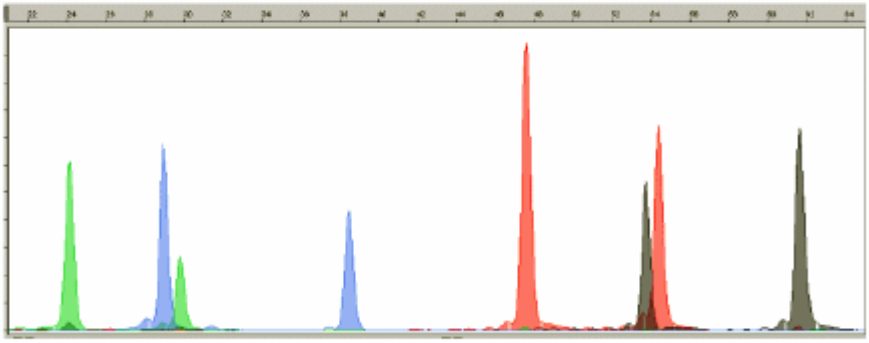


Figure 6. Overview of the SNaPshot™ Procedure (ABI, 2000)

The positive controls provided with the kit use six primers that are specific to the included control DNA. Figure 7 shows both a table providing the expected outcome of the primers and the resulting electropherogram produced by the positive control reaction. The expected peak size does not necessarily match-up with the expected fragment length

(due to the small size and nucleotide base incorporation) and thus should be considered relative.

Multiplex Control Primer Mix	Length of Final Products (nt)	Signal Color	Heterozygosity
20A primer	21	Green	Homozygote
28G/A primer	29	Blue/green	Heterozygote
36G primer	37	Blue	Homozygote
44T primer	45	Red	Homozygote
52C/T primer	53	Black/red	Heterozygote
60C primer	61	Black	Homozygote



Primer	20A	28G/A	36G	44T	52C/T	60C
Relative size of SNaPshot product*	24	29	39	46	53	62

*:As determined by sizing against GeneScan-120 Liz size standard

Figure 7. SNaPshot™ Positive Control Electropherogram (ABI, 2000)

Through the use of sequencing to compare different WSMV genomes, three specific SNPs were targeted as potential SNP sites. The three primers would provide effective discrimination of the three isolates in question. From these sites and another section of WSMV that proved to be reasonably conserved, three SNP primers and one diagnostic primer included as an internal control were created to be used in multiplex with the SNaPshot™ kit. Table 5 includes the names and complete sequences of the four SNaPshot™ primers.

Table 5. Genomic Sequences of Created SNaPshot™ Primers

Primer Name	Length	Sequence (5'→3')
SNP-1	45	TATGGCCTTTGTTGTGATCTTTGGTGCTGGAAACACTAAACCTCG
SNP-2	50	CGCATAATGGCTCGAAGTGATGGCTTCGCAGCTTTGTACATCGGTTCAAT
WSMV-40	40	TGAGTTGTTCAGTTGTGGCATATCTGTTGTCGATAAGTTC
WSMV-DIAG	55	TCACAACACGCCAAGTTCCTAGTGCTCTGTTCTCTCGGGCACGTTGTGTTGATTAT

5µL of SNaPshot™ Multiplex Ready Reaction Mix, 3µL of Zymo Column purified PCR products (about 10ng), 0.2µL of each of the four SNaPshot™ primers at 10mM, and 1.2µL of deionized water were placed in a PCR tube. This tube was then placed in a thermal cycler and amplified in a program advised by the SNaPshot™ kit protocol. Table 6 contains the specifics of the SNaPshot™ Thermal Cycler program.

Table 6. Thermal Cycler Program of SNaPshot™ PCR

ddGTP Addition		
25 cycles		
Denature	Anneal	ddGTP Addition
10 sec.	5 sec.	30 sec.
96°C	50°C	60°C

Once the cycling program was complete, the products were stored at 4°C. Remaining unincorporated ddGTP were destroyed using shrimp alkaline phosphatase as described in the following section.

Shrimp Alkaline Phosphatase Purification

Due to the use of fluorescently-linked ddNTPs in the SNaPshot™ method, a purification step is required to remove any unused fluorescence that would interfere with visualization of results. The SNaPShot™ Protocol suggests using Shrimp Alkaline Phosphatase (SAP); this hydrolyzing enzyme, when given the proper alkaline environment and enough time, can effectively dephosphorylate ddNTPs, allowing for efficient removal of any confounding luminescence.

Amplicons created through the SNaPshot™ process were mixed with 1µL of SAP (at a concentration of 1U/µL), 2µL of 10X SAP Reaction Buffer, and enough nuclease-free water to bring the tube contents up to 20µL total volume. The entire mixture was then placed in a 37°C water bath for one hour to allow the SAP enough time to effectively dephosphorylate the sample. Immediately thereafter, the PCR tube was transferred to a 65°C heating block and left for 15 minutes; this step completely and irreversibly inactivates the SAP. The dephosphorylated sample was then stored at 4°C until it could be analyzed with the ABI 310 Genetic Analyzer.

CHAPTER III

RESULTS

Production of WSMV cDNA

All three isolates—Isolate 964, Plants A through D, and OSU—initially provided discouraging results on every sample that had been extracted, amplified, and analyzed using gel electrophoresis. Size standards were clearly visible (Lanes 1 and 4 in Figure 6 below), but the lanes associated with the given WSMV isolates were largely empty. When a lane would occasionally fluoresce, it would do so at a size that could be attributed to primer dimerization, far from the amplified 1300 nucleotide fragment that should exist. Figure 8 shows gel results product demonstrating the failed cDNA amplification.

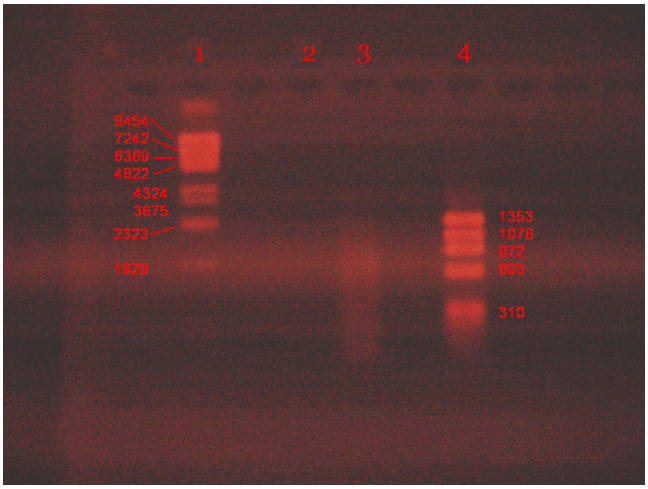


Figure 8. Electrophoresis Results of Initial cDNA Amplification. Lanes one and four are DNA size standards, Lane two is 964, and Lane three is Plant A.

Two different kits (one for RNA extraction and one for cDNA synthesis) had been used to prepare and reverse transcribe the viral genome, and electrophoresis could not provide any clue as to which had failed. Both the MagMAX™ (Ambion, Inc.) and SuperScript™ (Invitrogen, Co.) kits would have to be tested individually to investigate where the cDNA synthesis had failed.

SuperScript™ RT-PCR Verification

An additional set of WSMV-specific primers were received from the plant Disease Diagnostic Laboratory at OSU-Stillwater. These primers amplify a known region 540 nucleotides in length. Often used in plant pathology laboratories to specifically diagnose the presence of WSMV, the two primers, WSMV-H8369 and WSMV-C8908, are diagnostic primers that anneal to a very highly conserved sequence within the coat protein coding region of WSMV isolates. Table 7 includes the names, sizes, and sequences of the two diagnostic primers.

Table 7. Genomic Sequences of WSMV Diagnostic Primers

Primer Name	Length	Direction	Sequence (5'→3')
WSMV-H8369	26	Sense	CAATCTAATAACGTATCTGTCATGGC
WSMV-C8908	22	Anti-sense	CATAATGGCTCGAAGTGATGGC

To investigate the efficacy of the RT-PCR process, these primers were used with the SuperScript™ kit and with the purified RNA stock of the OSU isolate. Two additional SuperScript™ reactions were performed alongside this diagnostic check but both contained the original XC1 and XV1 primers synthesized using the published sequence

from French et al. (McNeil, 1996; Chenault, 1996; French, 2005). The first sample contained Isolate 964 RNA, and the second sample contained Plant A RNA. Figure 9 shows the results of that particular experiment.

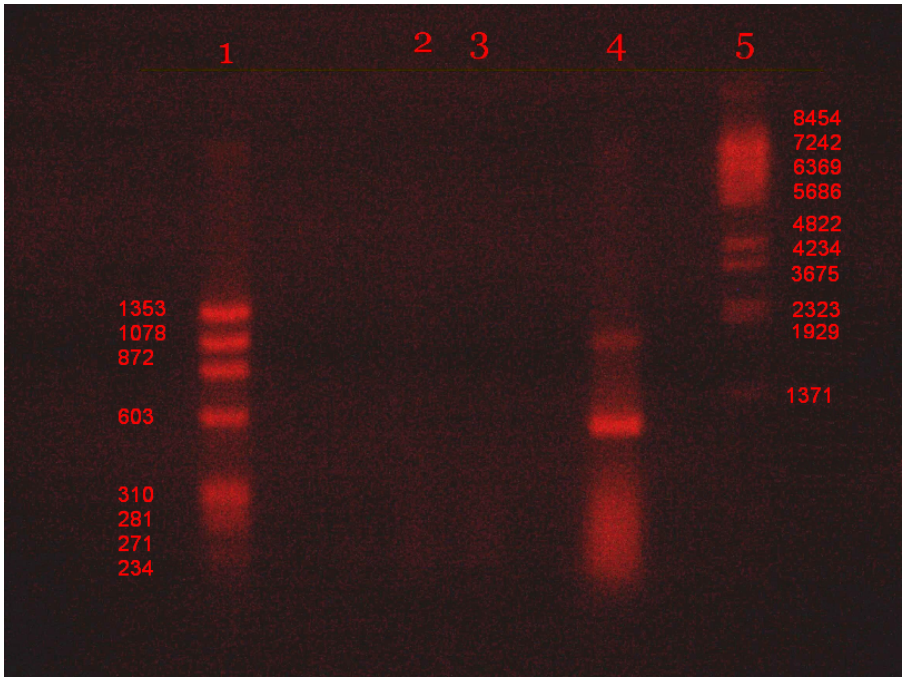


Figure 9. Electrophoresis Results of WSMV Diagnostic Primer Amplification. Lanes one and five are DNA size standards, lane two is Isolate 964, lane three is Plant A, and lane four is the OSU isolate.

The 550 nucleotide RT-PCR product seen in lane four verified two processes that were previously under question: the WSMV-H8369 and WSMV-C8908 primers directed the synthesis and amplification of the conserved region of the coat protein gene as expected when in the presence of WSMV cDNA. This result confirmed that the SuperScript™ kit, when given suitable template and primers, would indeed work properly. The proper template in this instance was viral RNA that had been extracted and purified elsewhere. These results narrowed the list of possible explanations for the failed RT-

PCR of our viral isolates to one of two additional problems: either the primers XC1 and XV1 used for RT-PCR were faulty or the MagMAX™ extraction protocol was failing.

MagMAX™ Extraction and Coat Protein Primer Verification

Based upon their direction of amplification within the cDNA, and reassured that the diagnostic primers properly amplify their target region, XV1 and XC1 were paired with WSMV-C8908 and WSMV-H8369 primers, respectively. Any failures in amplification would reflect negatively on the coat protein region primer and would warrant further inspection. As an additional check on the MagMAX™ extraction protocol, previously extracted RNA samples were subjected to SuperScript™ with the diagnostic primers and were included in the electrophoresis. Figure 10 shows the results of the experiment.

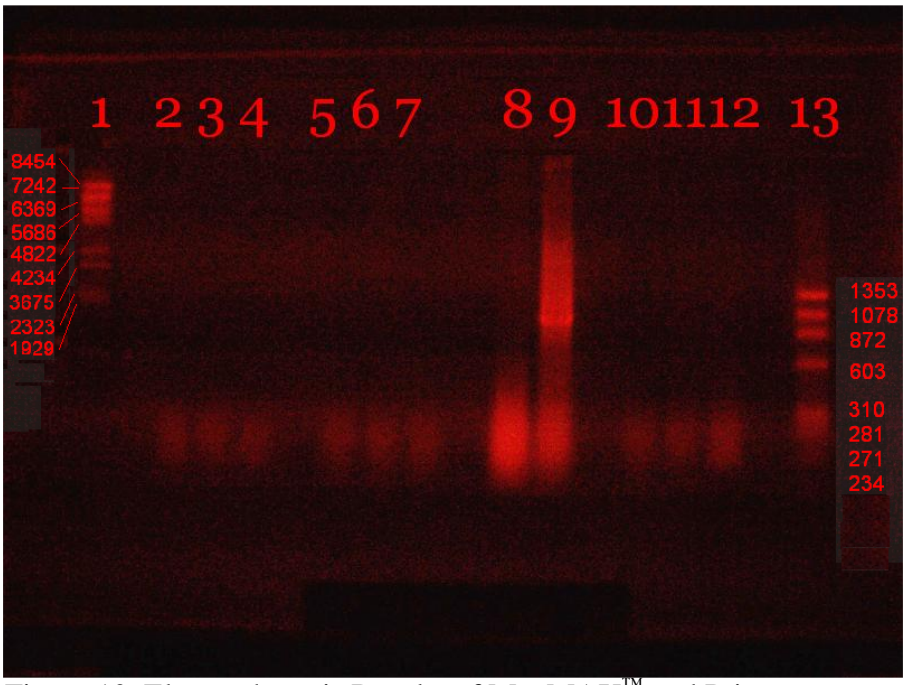


Figure 10. Electrophoresis Results of MagMAX™ and Primer Verification. Lane information can be found in Table 8 .

Table 8 lists the products associated with the given lanes listed in Figure 10.

Table 8. Lane Information of Figure 10

Lane Number	Sample Run	Primers Included	
		Sense	Antisense
1	DNA size standard–*BstEII*	N/A	
2	Plant A	WSMV-H8369	WSMV-C8908
3	Plant A	XV1	WSMV-C8908
4	Plant A	WSMV-H8369	XC1
5	Plant B	WSMV-H8369	WSMV-C8908
6	Plant B	XV1	WSMV-C8908
7	Plant B	WSMV-H8369	XC1
8	OSU	XV1	WSMV-C8908
9	OSU	WSMV-H8369	XC1
10	Control	WSMV-H8369	WSMV-C8908
11	Control	XV1	WSMV-C8908
12	Control	WSMV-H8369	XC1
13	DNA size standard–*HaeIII*	N/A	

The results of this verification experiment offered important clues to the viability of the protocols currently being used. Despite being paired with the diagnostic primers, Lanes 2 and 3 failed, affirming that no viral RNA was extracted from plant A using the MagMAX™ procedure. Lanes 8 and 9 were also of note. Lane 8, containing OSU RNA paired with XV1 and WSMV-C8908, failed to produce a properly amplified product. Lane 9, containing OSU RNA paired with WSMV-H8369 and XC1, produced a band just below 1,000 nucleotides. This result suggested that the XV1 primer might be failing and thereby responsible for the lack of amplified cDNA when previously paired with its antisense primer, XC1.

Modification of the MagMAX™ Extraction Procedure

Production of cDNA seemed to fail in the extraction phase despite following the given protocol explicitly. Thus, barring the possibility that the kit was somehow

defective, some individual step within the MagMAX™ protocol was possibly suspect for causing the entire method's failure. Due to its abbreviated interaction with the RNA sample, the viral lysis procedure was initially targeted as the limiting step, and so the lysis step was modified by extending the incubation time and increasing the temperature for lysis from room temperature to 65°C.

Once the 802µL of Lysing Buffer/Binding Solution were added to the liquid sample containing virus, the tube was incubated at 65°C for one hour instead of room temperature for four minutes to give the lysis buffer the greatest chance of releasing RNA into solution. This step completely replaced the four minute extraction period at room temperature directed by the MagMAX™ protocol.

Following extraction and magnetic bead interaction, an additional change was enacted to better allow the beads to elute their captured RNA into solution. Once the pre-warmed elution buffer was added to the magnetic bead solution, the samples were placed on a 65°C heating block for five minutes. This adjustment was meant to allow the beads ample time to completely elute their captured RNA. Following this modified elution step, the protocol completed as normal.

To compare the results of the new longer extraction method, an additional extraction method using Phenol/Chloroform extraction was also performed on the RNA. Both extraction methods were applied to all three WSMV isolates, and the resulting RNA was subjected to RT-PCR using primers WSMV-H8369 and XC1. If either extraction method succeeded, a single band just below 1,000 nucleotides would be expected. Figure 11 shows the results of this experiment.

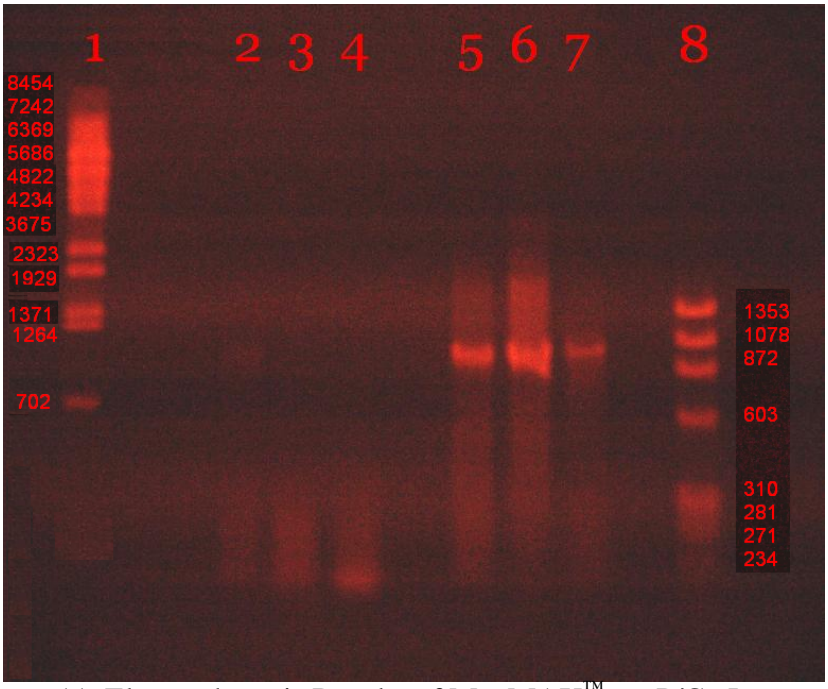


Figure 11. Electrophoresis Results of MagMAX™ vs. P/C. Lanes one and eight are DNA size standards. Lanes two through four are P/C of A, 964, and OSU, respectively. Lanes five through seven are Long Extraction MagMAX™ of A, 964, and OSU, respectively.

The longer extraction method used with the MagMAX™ kit resulted in better extraction of the RNA from the viral particles.

Fixing the Sense Primer, XV1

Since XV1's usefulness remained in doubt, a new sense primer was created to be used with XC1 in the RT-PCR procedure. This primer, named Carver-1 (abbreviated C1), annealed just downstream of XV1. While it did not amplify a region as large as XV1 when paired with XC1, it did capture the entire coat protein gene, amplifying from nucleotide 8117 to the end of XC1 at nucleotide 9372 (numbers published for the Sidney 81 genome from GenBank, Accession #AF057533). A new SuperScript™ reaction was performed with RNA extracted using the modified protocol for MagMAX™, with XC1 and C1 as primers. A successful reaction would amplify the nearly 1,300 nucleotide

region between the two primers and exhibit a single band at that point. As shown in Figure 12, C1 in conjunction with XC1 directs the amplification of a product of the expected size.

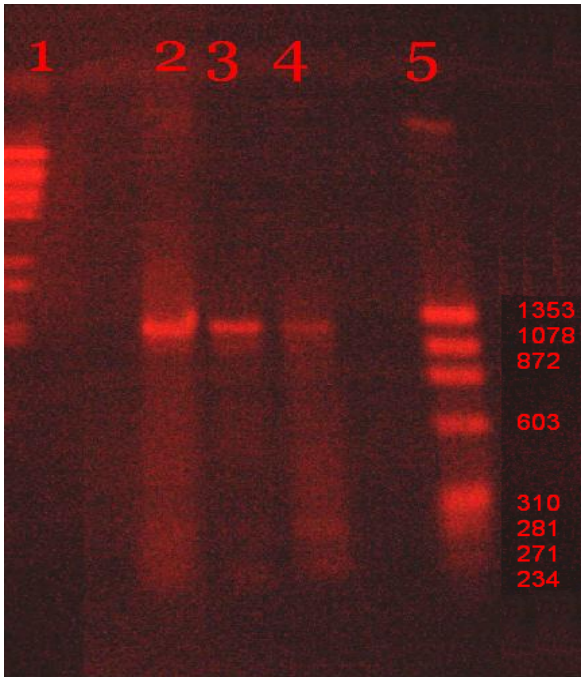


Figure 12. Electrophoresis results of XC1/C1. Lanes one and five are DNA size standards. Lanes two through four are A, 964, and OSU, respectively.

Considering the success of this experiment, all further SuperScript™ RT-PCR reactions were done with C1 and XC1 as the sense and antisense primers, respectively.

Finding Adequate Restriction Sites

Both *Hpa*II and *Ava*I were used as tests meant to determine the feasibility of applying the methods of AFLP to discriminate among WSMV isolates. Proper site recognition and subsequent cutting could provide a new method for effective discrimination. Figure 13 shows a graphical display of how the cDNA should react to the two restriction endonucleases.

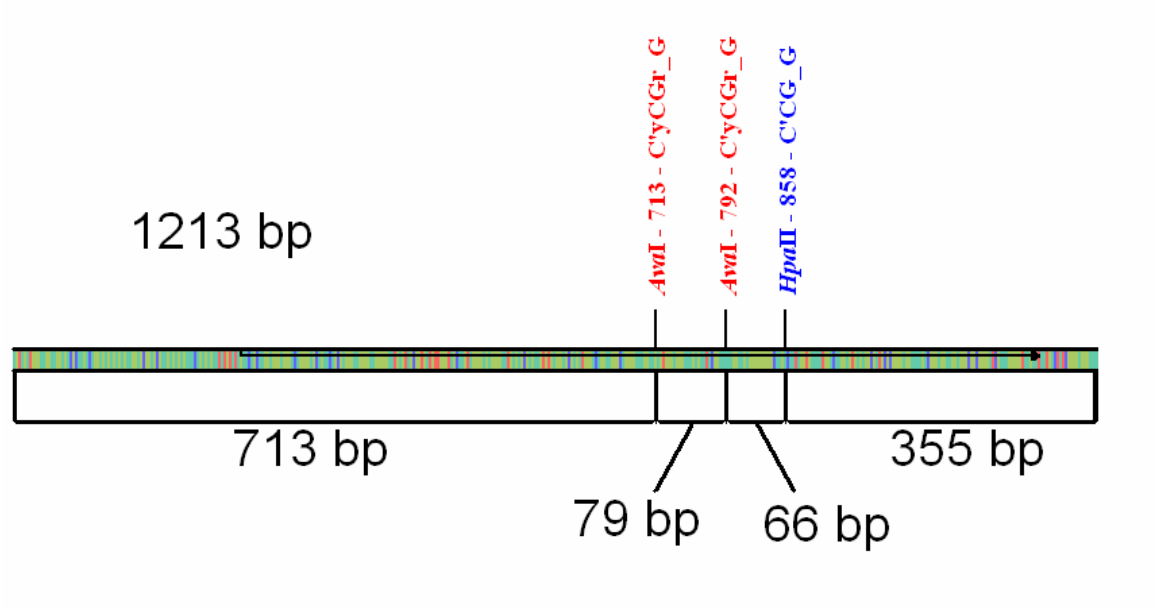


Figure 13. Restriction Sites for *Hpa*II and *Ava*I in WSMV

After completing the two separate incubation periods required for proper DNA cleavage, the resulting samples were loaded on a gel and electrophoresed. Figure 14 shows the gel after 200µL of EtBr were used as a stain. The gel is seen above a UV light source.

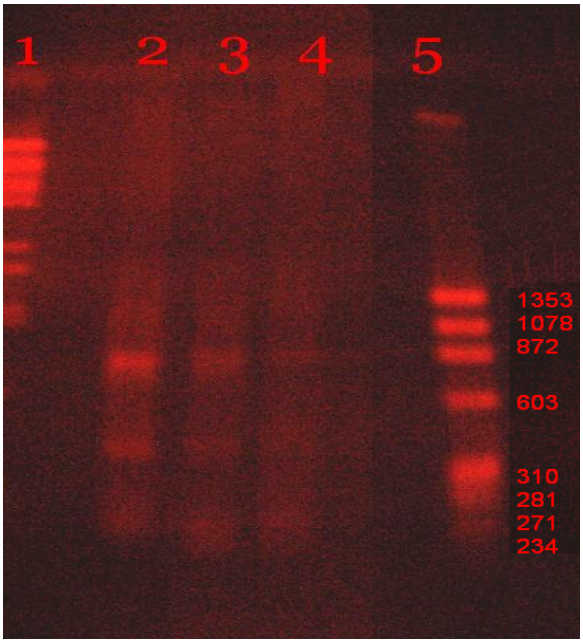


Figure 14. Electrophoresis Results of AFLP. Lanes one and five are DNA size standards. Lanes two through four are A, 964, and OSU, respectively.

Four products should have resulted from the three separate cuts to the cDNA. Due to the low resolution power of agarose gel, no distinction can be made between the 66 and 79 nucleotide bands. However, the three bands that are seen correlate with the expected results; the highest band is roughly 700 nucleotides, the middle band is roughly 350 nucleotides, and the lowest band is roughly 75 nucleotides. While these results are encouraging in that they did exactly as expected, they are also quite discouraging as all three isolates produced the same restriction profile. Discrimination of WSMV isolates is, of course, the driving goal behind these experiments.

A complete restriction map was created through the use of pDRAW32, a freeware program available from AcaClone. A thorough inspection of this restriction map (Appendix A) ultimately proved the futility of using AFLP methods on this viral RNA genome for a number of reasons:

- AFLP works well on larger genomes in which a number of fragments will be created by restriction digestion. However, a DNA molecule only 1213 nucleotides long is limited in its overall endonuclease sequence recognition capability.
- While many restriction endonuclease recognition sites do exist within WSMV's coat protein coding region, many that do are common to most isolates and therefore lack discrimination capability.
- Unfortunately, those restriction endonucleases that are both specific and potentially discriminating are not commercially available.

Shortly after this experiment, AFLP techniques were abandoned in favor of methods that were more WSMV (and viral genome) specific.

Sequencing One-Quarter of the Genome

An attempt to apply the Sanger sequencing method to one nucleotide residue (guanine) initially proved fruitful. The resulting electropherogram should have the characteristic appearance of a Sanger-based sequence with initial fluorescence spikes that are high, but relative fluorescence units (rfus) of each subsequent peak would dissipate. Each dGTP that attached to the growing nucleic acid would create a higher statistical probability that a chain-ending ddGTP would be added instead, thus more products would be created earlier in the reaction. The resulting ddGTP-terminated chains could be separated by capillary electrophoresis and discriminated by color.

The initial desire to find only the guanine residues was due to an observation that occurred during the analysis of the coat protein region sequencing data. The larger

majority of single nucleotide substitutions of fixed SNPs (or changes that occurred complementarily in both the 5' and 3' sequenced strand) that differentiated a WSMV isolate from the consensus sequence were cytosine/guanine residue changes. As a result, greater variability and thus greater discrimination could be exploited by examining one of those two nucleotides exclusively.

Using a Labeled Primer

Ostensibly, the first test of single nucleotide sequencing appeared to provide Sanger-esque results. However, far too many ddGTP-ended chains were visible in the electropherogram (suggesting a genome composed primarily of guanines), but the large peaks initially with an overall, per-base decline in fluorescence does mimic the expected results. Unfortunately, the non-linear deterioration which appears rather prominently at 20 nucleotides detracts from the potentially positive experimental results. Figure 15 shows the original test using a FAM-linked Carver-1 primer with a mixture containing 5% unlabeled ddGTPs.

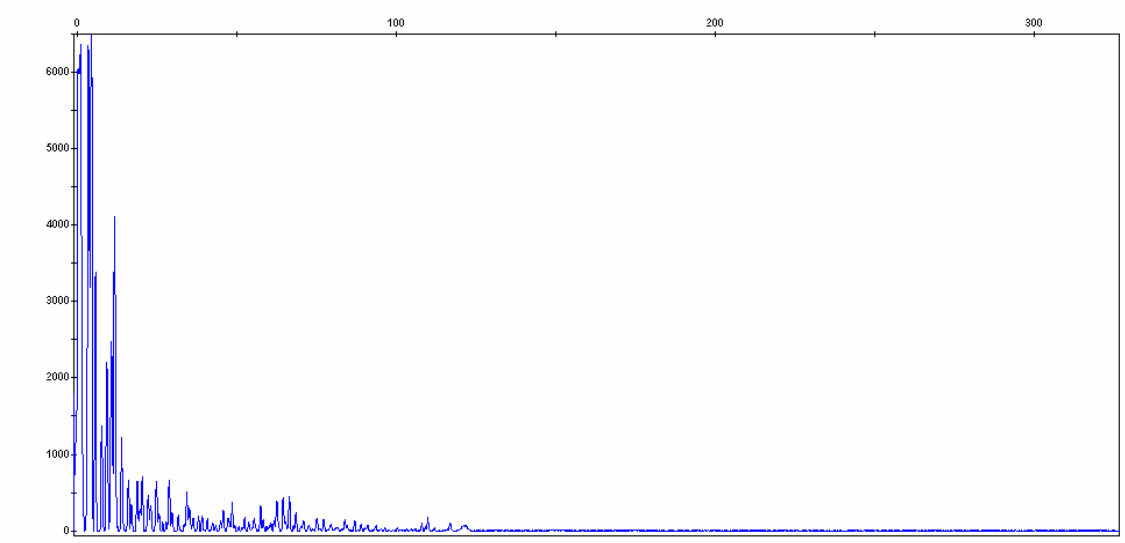


Figure 15. Initial Single Nucleotide Sequencing Electropherogram

While not a perfect representation of what a sequencing method in four parts would display, the results of this test were encouraging, and further investigation was merited. The same experiment was performed on varying isolates multiple times. However, no electropherogram produced after the initial test would provide reproducible results. In fact, the subsequent tests generally yielded an electropherogram that defied statistical logic and Sanger methodology by having substantial peaks appear far beyond the normal linear decline that should exist. Figure 16 shows a typical result created during successive tests.

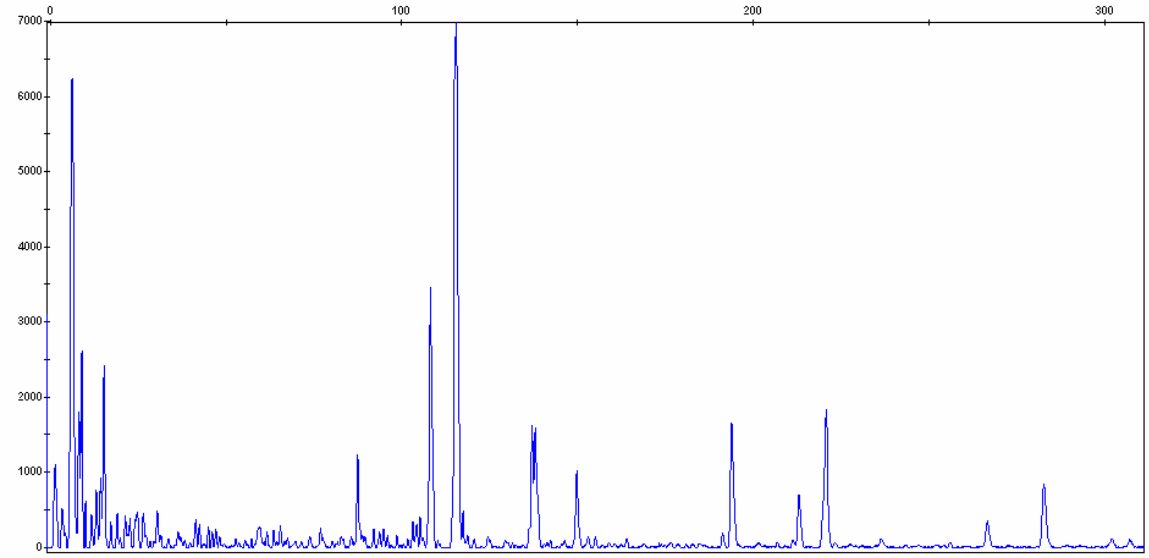


Figure 16. Subsequent Single Nucleotide Sequencing Electropherogram

Using a Labeled ddGTP

A major question associated with single nucleotide sequencing had yet to be addressed at this point: were the nucleic acid chains terminating early due to the addition of a ddGTP or were they simply terminating early for other reasons? The use of a FAM-labeled primer would not allow the distinction between these two possibilities; any

interruption, no matter the cause, would appear the same in the test's resulting electropherogram. Thus, the alternate approach was taken into account for the potential ambiguity inherent to the current system.

An unlabeled primer will not glow unless a fluorescently-linked label is attached to it. If the only label available to label terminated chains is ddGTP that has been linked to a fluorescein molecule, then the nucleic acid created from the primer will not fluoresce unless a ddGTP has capped and terminated the growing DNA chain. After successive rounds of PCR, any free-floating fluorescent ddGTP can then be dephosphorylated. In theory, the associated mixtures of labeled ddGTP with unlabeled dGTP would produce an electropherogram much like an expanded version of Figure 13 with blue peaks corresponding to each guanine residue within the genome.

Such an experiment was performed using amplicons from all three isolates as the template cDNA. Initially, the results seemed encouraging as each isolate produced electropherograms that not only exhibited a wide range of peaks, but did so over the course of several hundred nucleotides. Closer inspection, however, showed that Plant C (Figure B-1, Appendix B), Isolate 964 (Figure B-2, Appendix B), and OSU (Figure B-3, Appendix B) all exhibited exactly the same pattern of peaks on all electropherograms. This pattern, likely due to baseline peaks attributed to a lack of purity within the sample, was consistent across every attempt at single nucleotide sequencing with an unlabeled Carver-1 primer and fluorescein-labeled ddGTP.

Sequencing the Entire Genome

When the single nucleotide sequencing method failed to produce any viable results, the BigDye® Terminator Kit was used to sequence the entire genome. With the ABI 310 Genetic Analyzer's ability to separate and display the different colored dyes in different histograms, the goals of the original one-quarter sequencing protocol of finding individual guanine differences could still be accomplished by viewing only the guanine residues. All three isolates were subjected to the BigDye® Terminator Kit sequencing protocol as previously described and the red, green, yellow, and orange (which corresponds to the LIZ 120 internal sizing standard) dyes were filtered out to leave only the blue electropherograms.

Initially, the additional Genetic Analyzer instructions in the BigDye® Terminator Kit went unheeded and the samples were run as any fragment analysis sample would be; the resulting graphs were nonsensical if they produced sizing data at all. Even after the proper Run Module (P4StdSeq (1 mL) E), Mobility File (DT310POP4{BD}v2.mob), and Matrix Standards (310/377 BigDye® Terminator v1.1 Matrix Standards) were applied, the results did not improve. When the electropherograms of the three isolates were compared together, only two minor peaks exhibited any differences; unfortunately, their differences were deemed superficial as the corresponding peaks did exist in the other graphs, but they simply missed the minimal fluorescence required to be labeled as a peak.

Plant C (Figure C-1, Appendix C), Isolate 964 (Figure C-2, Appendix C), and OSU (Figure C-3, Appendix C) produced complex histograms with largely identical peaks. None of the peaks reached above 1600 rfus and the majority of the peaks did not reach above 1000 rfus. This lack of fluorescence and the similarity in the

electropherograms can likely be attributed to a high baseline more than actual sample dissimilarities. Due to the habitual failures of the sequencing protocol and the associated lack of troubleshooting knowledge, sequencing was abandoned after the early successes of SNP-specific analysis.

Purifying Samples Following Fluorescent Attachment

Putting PCR samples through a purification protocol after the incorporation of fluorescent labels is invaluable to avoid complicated and sometimes uninterpretable results. If the samples are not purified, baselines will appear higher than normal, peaks will occur with less sharpness, and unexpected results will manifest, all of which detract from any potentially positive interpretations. Of course, the type of purification protocol depends solely upon the type of fluorescent protocol performed in the previous methods. Two different purification protocols were used on differing sample types: SAP was used on any SNaPshot™ sample, and Centri-SEP Column Purification was used on any sequencing-based sample.

The results before and after purification were markedly different. SNaPshot™ samples produce results with a single, strong, uncontaminated peak for every SNP primer used in the multiplex kit. Figures 17 shows a single-primer SNaPshot™ result without an additional fluorescent purification step.

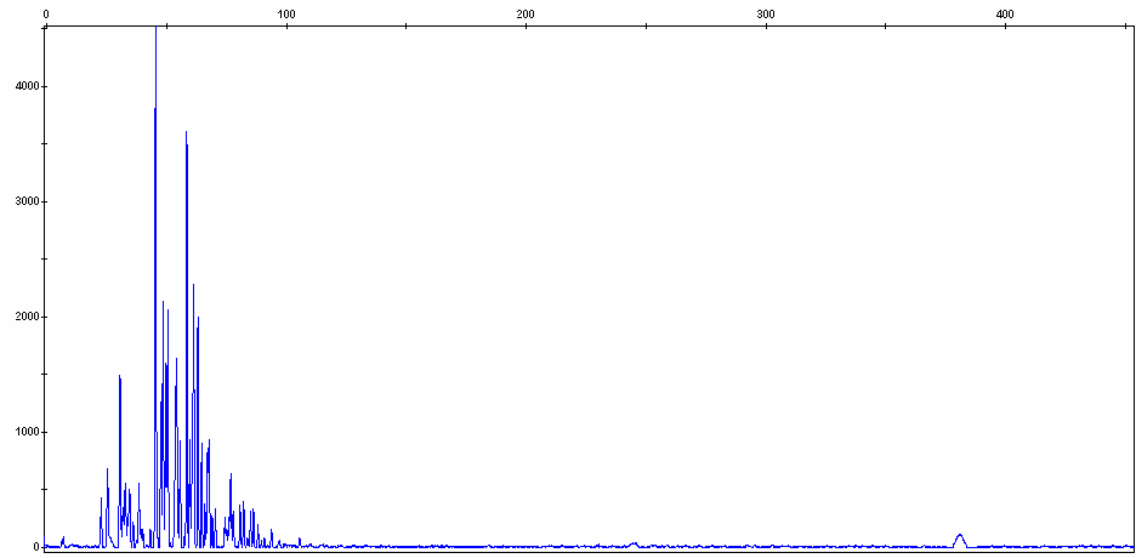


Figure 17. Single-Primer SNaPshot™ Electropherogram before SAP Purification

Figure 18, on the other hand, is the same sample after it has been purified by SAP.

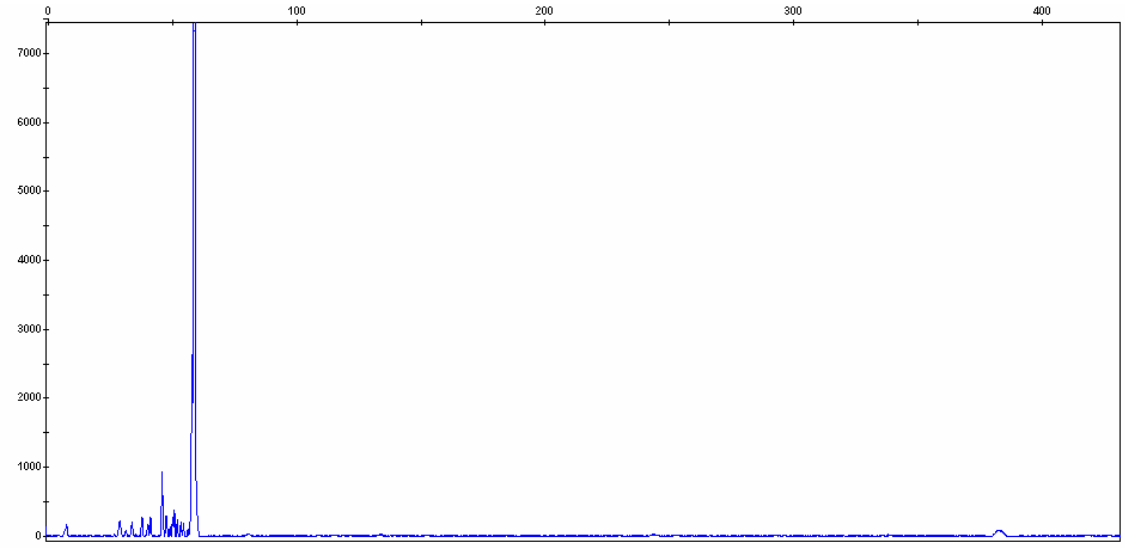


Figure 18. Single-Primer SNaPshot™ Electropherogram after SAP Purification

The same situation occurred when a sequenced sample was not subjected to the Centri-SEP Column Purification after fluorescent labeling. In fact, such a sample analyzed before being Centri-SEP purified would provide no sizing data at all, thus ultimately showing the importance of a cleaning step after adding any fluorescently-linked molecule to an analyzed product.

Finding Isolate Differences Using SNaPshot

The nucleotide sequences of the coat protein genes of the three strains of WSMV used in this study revealed several single nucleotide polymorphisms. SNaPshot™ primers terminating one nucleotide 5' to these SNPs were synthesized (Invitrogen, Carlsbad, CA) and used with coat protein cDNA as template for the SNaPshot™ reaction. Our goal was to incorporate all SNP primers in a single multiplexed SNaPshot™ reaction that would discriminate the three WSMV strains.

Detection of the allelic form of each polymorphic site is accomplished by electrophoretic separation of each primer (extended one nucleotide by a dideoxynucleotide) and determination of the color emitted by the fluorescent tag. The expected genotype of each strain was known from the sequence data produced for each isolate. Table 9 identifies the dideoxynucleotide and color that were expected for each WSMV isolate with each SNP primer. The WSMV-DIAG primer targets a site in the WSMV genome that is conserved among all strains of the virus. This primer was added to the multiplex as an internal control primer for WSMV.

Table 9. Expected Results of the SNaPshot™ Multiplex Reaction

Primer	Length	Isolate					
		Plant C		Isolate 964		OSU	
		Base	Color	Base	Color	Base	Color
WSMV-40	40	C	Black	A	Green	C	Black
SNP-1	45	T	Red	C	Black	T	Red
SNP-2	50	T	Red	T	Red	C	Black
WSMV-DIAG	55	G	Blue	G	Blue	G	Blue

A SNaPshot™ electropherogram should consist of four fluorescent PCR products at 41, 46, 51, and 56 nucleotides (the length of the primer plus one additional base) of the expected color. As shown in Figure 7 on page 30, primer length is not the sole determining factor of a small DNA fragment's apparent size. Since the primer sizes are relatively small, the type of ddNTP incorporated to the elongating DNA chain can shift the peaks, making them appear larger than they actually are. Thus, amplicon size estimates should be considered relative and may vary from expectation depending on which nucleotide is incorporated at a particular site.

The results for Plant C proved encouraging. Four well resolved peaks of the expected colors were visualized at sizes that were separated by roughly five nucleotides. Figure 19 shows the electropherogram produced by multiplexing all four primers with Plant C as the cDNA template.

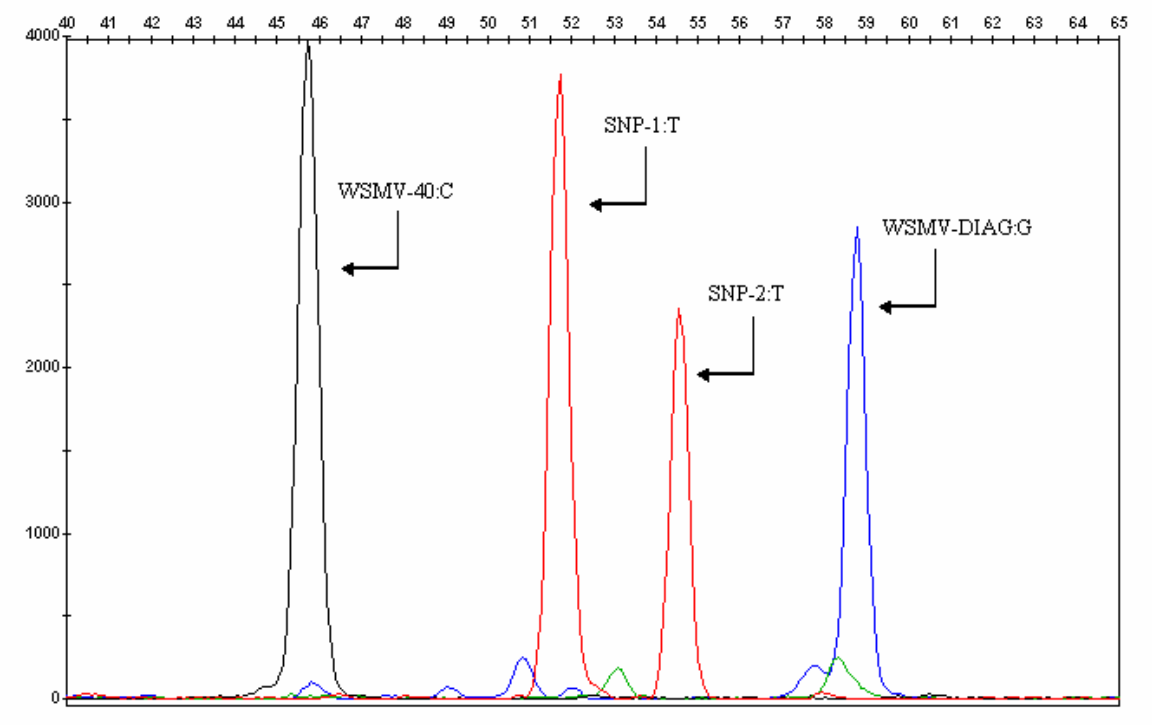


Figure 19. SNaPshot™ Electropherogram of Plant C

While still encouraging, Isolate 964 provided results that were more complex than the SNaPshot™ electropherogram of Plant C. Every expected peak exists within the Isolate 964 electropherogram, however several additional, unanticipated products were visualized. For two of the three SNaPshot™ sites however, the expected genotype is actually secondary to a larger, more prominent peak. In fact, the only SNP visualized in the results produced for Isolate 964 that is a single genotype is associated with the WSMV-DIAG diagnostic primer. Figure 20 is the electropherogram produced by the Multiplex SNaPshot™ Protocol when Isolate 964 is included as cDNA template.

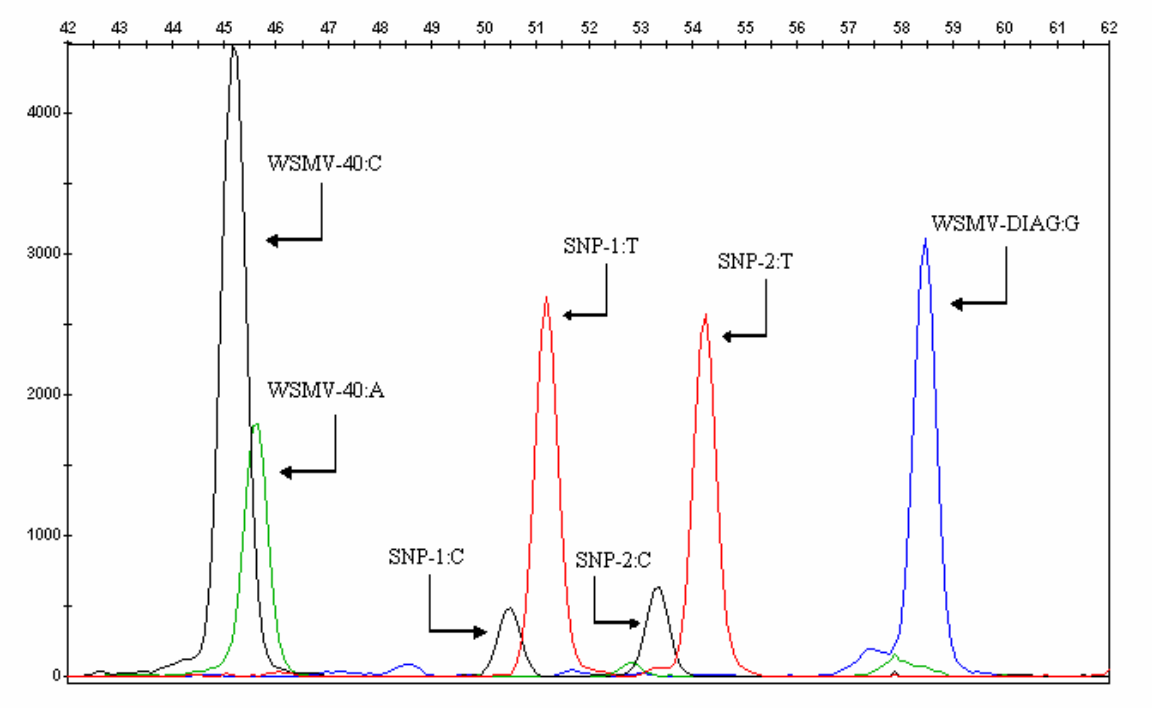


Figure 20. SNaPshot™ Electropherogram of Isolate 964

Possible interpretations for these results are that either there is some technical artifact associated with the SNaPshot™ reaction, or that more than one population of WSMV (of differing genotype) is present in Isolate 964. Inasmuch as the Plant C isolate did not exhibit a SNaPshot™ profile suggestive of a mixed genotype, the 964 WSMV isolate consists of more than one viral genotype indicating a mixed infection of the wheat plant from which Isolate 964 was originally obtained. One strain contains a T in the genome at the site targeted by the WSMV-40 primer (as evidenced by the incorporation of ddATP at that site). The other strain harbors a G at this position (as evidenced by the incorporation of ddCTP). Based upon the relative fluorescence of ddATP and ddCTP incorporated at the WSMV-40 site, the relative proportions of the ddATP and ddCTP genotypes are 29% and 71%, respectively. Similar proportions for the different

genotypes revealed by the SNP-1 and SNP-2 primers were also seen. Mixed infections of WSMV have been documented before (Stenger, 2002).

The OSU isolate also suggested the presence of more than one WSMV strain. The WSMV-40 SNP site consisted of two genotypes, T and G, as was observed for Isolate 964. However, the proportion of the T-type strain was reduced from 29% to 17% (Figure 21, Table 11). In contrast to Isolate 964, the SNP-1 polymorphism consisted entirely of genotype A. The SNP-2 polymorphism again demonstrated a mixed genotype, but the proportion of type A and type G virus in the mixture were roughly equal (Table 11, page 63). Given the results for the WSMV-40 polymorphism, there must be a minimum of 3 different strains of WSMV present in the OSU isolate to account for the SNaPshot™ results. If a maximum of two populations existed within the Isolate 964 sample, peak ratios would occur in only two percentages. The WSMV-DIAG primer produced a single genotype confirming all isolates are indeed WSMV.

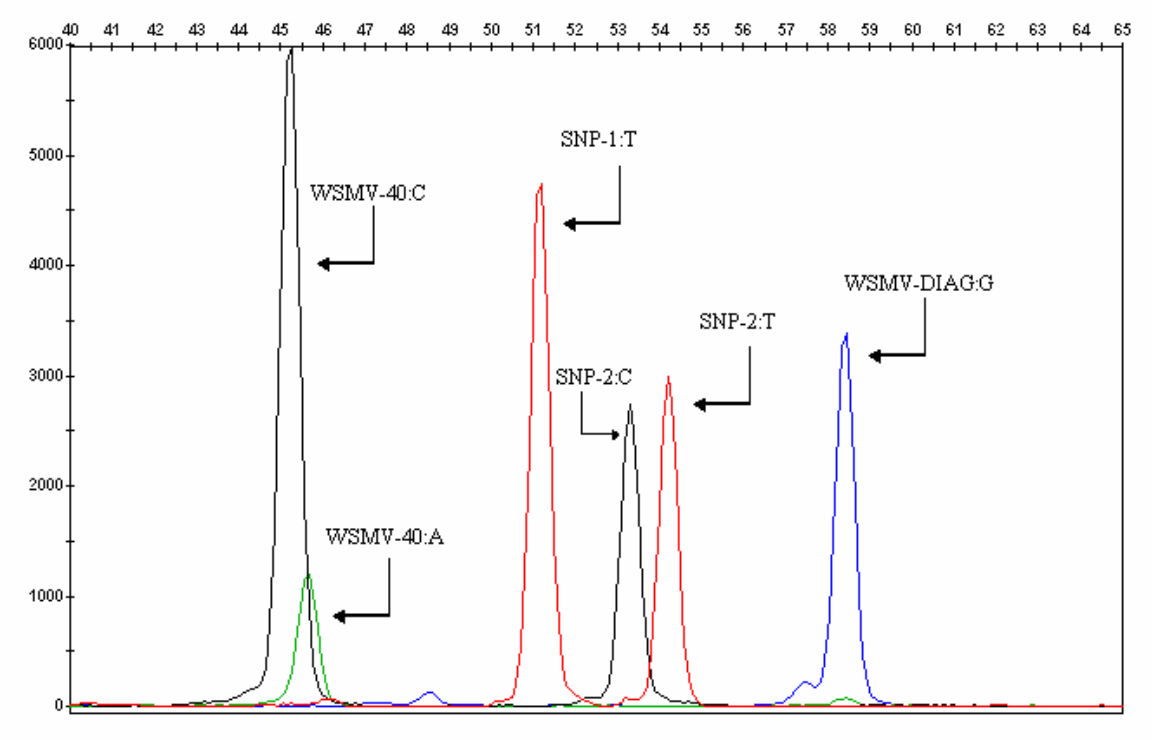


Figure 21. SNaPshot™ Electropherogram of OSU

While all three isolates were qualitatively discriminated by the multiplex SNaPshot™ assay, the quantitative differences in peak heights at heterozygous positions in the coat protein gene increases the discriminatory power of the assay.

Reproducibility of SNaPshot Results

Several additional results further support the conclusion that the 964 and OSU isolates of WSMV consist of more than one strain of the virus. First, if the SNaPshot™ reaction is performed without added cDNA template reverse transcribed from viral mRNA, no fluorescent signal is detected anywhere in the electropherogram. Thus, under SNaPshot™ conditions, interactions among the primers that could account for the fluorescent products seen do not occur. Figure 22 shows an electropherogram produced

from a SNaPshot™ reaction that contained all reagents except it lacked any cDNA template. The electropherogram shows nothing more than a baseline reading that never reaches above 115 rfus.

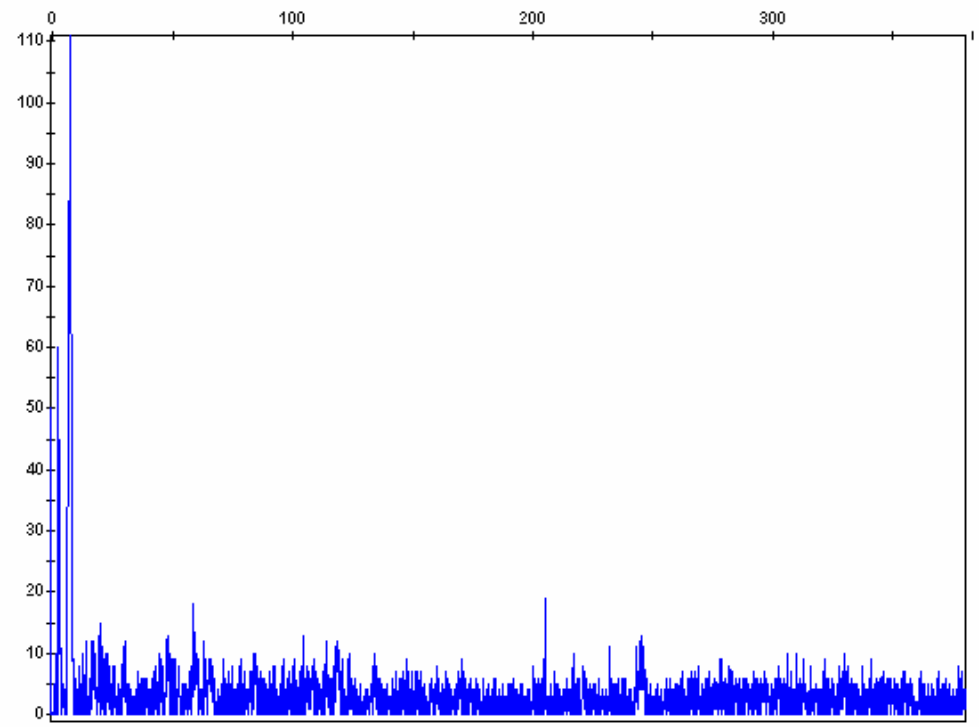


Figure 22. Electropherogram Produced by a Blank SNaPshot™ Reaction

Secondly, if the apparent heterozygous positions in the coat protein gene are the result of artifacts introduced by the SNaPshot™ reaction, they are highly reproducible as repeated extracts of mRNA from the different WSMV isolates give virtually identical results in terms of the relative proportions of the different genotypes present in each isolate (Table 11). One would not expect such a high degree of reproducibility if the heterozygous sites in the genome were the result of an artifact. Appendix D contains the electropherograms of additional SNaPshot™ retests of Plant C (Figures D-1 & D-2), Isolate 964 (Figures D-3 & D-4), and OSU (Figure D-5 & D-6).

Negative Control of the SNaPshot Reaction

Due to the relatively small size of the primers used in the Multiplex SNaPshot™ reaction, the question arose as to whether or not the SNaPshot™ primers would hybridize to genomic targets in DNA from organisms other than WSMV. Maximal utility of the multiplex reaction will be realized only if the assay is WSMV-specific.

To test for spurious binding, all four primers used in the SNaPshot™ multiplex reaction were entered as nucleotide search criteria into the nucleotide search function of GenBank's nucleic acid database. Any extreme deviation (>15%) from the primer sequence would be preferentially avoided in favor of the WSMV template. Matches to published WSMV sequences present in GenBank were removed to prevent confusion, and Table 10 presents the remaining matches by percent fit to the given primer. Given the results in Table 10, it is unlikely the SNaPshot™ products seen in this assay are derived from any organism other than WSMV.

Table 10. GenBank Nucleotide Search Matches for the Synthesized Primers

Primer	Matching Species	Number of Matches	Percent Complementarity to Primer
WSMV-40	*Mus musculus*	5	57, 55, 47, 45, 45
	Homo sapiens	4	57, 57, 57, 55
	Pan troglodytes	1	57
	Saccharophagus degredans	1	47
SNP-1	*Mus musculus*	7	42, 42, 42, 42, 42, 42, 42
	Oryza sativa	2	88, 51
	Angiopteris evecta	1	42
	Oreochromis miloticus	1	46
	Rattus norvegicus	1	42
SNP-2	*Bacillus pumilus*	1	58
	Pseudoalteromonas haloplanktis	1	52
	Serratia proteamaculans	1	42
	Ajellomyces capsulatus	1	36
	Methylibium petroeiphilum	1	34
	Mus musculus	1	34
	Escherichia coli	1	34
	Anopheles gambiae	1	34
	Candida glabrata	1	34
WSMV-DIAG	N/A	N/A	N/A

An additional approach was to actually assay a different viral genome in the SNaPshot™ assay to see if any product was produced. Three isolate clones of Barley Stripe Mosaic Virus (BSMV) were subjected to the SNaPshot™ assay. The three samples were capillary electrophoresed with the ABI 310 Genetic Analyzer. The electropherograms produced by SNaPshot™ analysis of BSMV Alpha (Figure E-1, Appendix E), BSMV Beta (Figure E-2, Appendix E), and BSMV Gamma (Figure E-3, Appendix E) contained only one peak above 300 rfus and no peaks above 150 rfus in the area of interest between 40 and 60 nucleotides.

CHAPTER IV

DISCUSSION

The threat of agroterrorism has steadily increased as more and more weaknesses have been identified within the world's agricultural economy. To aid in our preparedness, this research attempted to funnel the combined knowledge of agriculture, plant pathology, and forensic science into a new methodology that could support all the involved fields. By using WSMV as a model system, an assay was developed to aid extension agents and diagnostic laboratory scientists in discriminating differing isolates of a virus. When applied to an outbreak thought to be intentional, this assay could assist law enforcement personnel in tracing, and ultimately identifying the original clandestine laboratory where the virus was cultivated before the outbreak.

The SNaPshot™ assay effectively discriminated three WSMV isolates qualitatively. While this qualitative distinction among isolates is important, an equally important capability of the assay is the potential for discriminating isolates consisting of mixed viral strains based also on their quantitative differences as revealed by the SNaPshot™ assay. In a 2002 paper for *Virology*, Drake Stenger noted that per isolate discrimination existed due to the "multiple and very closely related genotypic variants [that] are resident within a single isolate." In other words, a wheat plant, while infected by the same isolate throughout, may contain a varying proportion of similar, but slightly

different viral particles. This assay has managed to exploit and visualize such differences and apply them for forensic purposes using technology readily available in most current DNA typing laboratories.

Current WSMV Diagnosis Methods

The start-of-the-art technique in confirmatory testing for WSMV requires several time-intensive steps. The scientist utilizes a set of WSMV-specific primers that bind to a highly-conserved region within the virus's coat protein. After completing PCR, samples of the amplicons are mixed with a running dye and electrophoresed through an agarose gel for a predetermined amount of time. The resulting gel is stained and, if WSMV is present, a specific band of expected size (roughly 500 nucleotides in length for the WSMV-H8369 and WSMV-C8908 primers) will fluoresce under UV light. While effective, this methodology and the technology required to complete it are rapidly becoming outdated.

The SNaPshot™ assay described in this paper is an advancement in both the efficiency and the information gained when compared to the currently used diagnostic system. While incorporating the original WSMV target sequence from the current diagnosis method, the addition of other SNP analysis primers makes the proposed system more robust and informative, especially for forensic attribution. The electropherogram produced from a sample will only fluoresce at the 55-nucleotide length primer location when WSMV is present. However, genetic characteristics of the strain(s) under study and a more thorough understanding of the specific isolate being analyzed can be obtained at the same time, making this system far more useful for forensic identification. The

current diagnostic test can only reveal the presence of WSMV as the causative agent for a plant infection, whereas the multiplex SNaPshot™ assay, with the added diagnostic primer, has the potential to identify the pathogen and trace its origin.

SNaPshot Result Analysis

Initially, the SNaPshot™ electropherograms inspired more confusion than insight. Only Plant C produced an electropherogram that matched the expectations of the sequencing data. The 964 and OSU isolates produced electropherograms that were complicated by additional PCR products at target priming sites suggesting heterozygosity in genotype. For the 964 and OSU isolates, multiple genotypes at the WSMV-40 and SNP-2 sites were detected (Table 11). At the WSMV-40 site, two peaks were encountered in Isolate 964. In a subpopulation of the virus, about 29% of the total had a T residue at the site whereas 71% exhibited G. Transversion (purine to pyrimidine or pyrimidine to purine) polymorphism substitutions, although occurring, are generally uncommon and are heavily outnumbered by transition (purine to purine or pyrimidine to pyrimidine) polymorphisms (Decker-Walters, 2004).

In contrast to Isolate 964, the OSU strain, which also exhibited either T or G residues at the WSMV-40 site, exhibited much different proportions of the two subpopulations with about 17% expressing a T residue and 83% expressing a G residue. Similar results were seen with Isolate 964 and OSU when considering the polymorphism revealed with the SNP-2 primer; the two isolates exhibit the same polymorphism but with vastly differing proportions of the contributing strains of WSMV (Table 11).

The reproducibility of the SNP profiles from sample to sample and extract to extract strongly suggested that the profiles are real and not the result of an artifact associated with the SNaPshot™ kit. Likewise, area ratios between peaks connected to the same primer were extremely similar on successive tests of the same isolate. Table 11 shows the ratio of area of the smaller peak over the larger peak at a primer location containing two peaks.

Table 11. Percent Total Area Ratio Per Genotype of Heterozygous Locations

Isolate (Run)	Primer		
	WSMV-40	SNP-1	SNP-2
964 (1)	28.87% A	15.05% C	19.81% C
	71.13% C	84.95% T	80.19% T
964 (2)	29.74% A	15.38% C	19.81% C
	70.26% C	84.62% T	80.19% T
964 (3)	29.19% A	15.12% C	19.73% C
	70.81% C	84.88% T	80.27% T
OSU (1)	17.07% A		48.82% C
	82.93% C		51.18% T
OSU (2)	17.66% A		48.75% C
	82.34% C		51.25% T
OSU (3)	16.96% A		49.13% C
	83.04% C		50.87% T

Table 11 shows every case where heterozygosity existed from three SNaPshot™ analyses of the 964 and OSU strains. The peak areas of each polymorphism were used to determine their contribution to the total peak area at that locus. Those percentages were compiled in the table for better comparison. Admittedly, the number of repeats remains small; however, a pattern of relative ratios has already begun to develop. The greatest difference between percentages occurs between the first Isolate 964 run and the second

Isolate 964 run at 0.87%. Remarkably, the third run of Isolate 964 falls between the first two despite being the product of a different extraction from new green leafy material.

It is interesting to note that Isolate 964 and OSU were harvested from crops that had suffered a natural outbreak of wheat streak disease and were therefore subject to the additional external forces in the environment. The only electropherogram free of marked heterozygosity produced from the three isolates came from the viral sample believed to contain a single strain of WSMV, Plant C. Its laboratory controlled infection and passage provided it with less chance to mix multiple viral strains of WSMV from different infection sites (perhaps in the wheat curl mites that transmit the infection during feeding) into the mixed strain infections that led to the source of Isolate 964 and OSU.

Use of the SNaPshot Assay for Forensic Purposes

One of the most important considerations throughout the course of this research was how feasible it would be for the SNaPshot™ assay to be performed in a crime laboratory. All of the steps and protocols considered had to be based on the technology and training readily available to scientists in modern forensics laboratories. Any additional equipment purchases or training would hamper the assay from widespread dissemination throughout the forensic community.

None of the protocols used during the assay require any additional purchases beyond the kit, matrix, and consumable reagent costs. Likewise, similar methods to the assay are already employed by forensic scientists on a regular basis; the same magnetic bead technology is applicable to both MagMAX™ and DNA IQ™ System extractions, just as PCR amplification methods are nearly identical for both SuperScript™ and

Indentifiler©/Profiler Plus© STR typing. This common methodology allows the multiplex SNaPshot™ assay to have a greater appeal, as minimal changes would be required to incorporate it into the daily workings of a crime laboratory.

The assay seamlessly integrates into a forensic laboratory because its approach to discrimination mirrors those utilized by modern DNA testing facilities. Simple, seemingly insignificant changes within the genome are tested to qualitatively compare samples of unknown origin with reference samples. Discrimination occurs when the resulting electropherograms of the tested samples can be compared for divergence from the consensus sequence. Those samples of similar origin will produce similar electropherograms and allow scientists to determine with relative scientific certainty that they were at one point recently part of the same commingled viral isolate.

Diagnosis and Discriminatory Power

Due to the combination of SNP-specific primers with the diagnostic primer, a single assay can diagnose and distinguish differing viral isolates. As shown in the results, the assay has the ability to do so with at least the first three strains tested via this protocol. The test in its current form has the ability to qualitatively distinguish many additional isolates. Although unlikely, if all available permutations of the three SNP analysis peaks were available in nature within WSMV, the assay would be able to distinguish 64 different strains. Any additional SNP-specific primers added to the multiplex kit would increase the discriminatory power of the test even further.

If quantitative differences observed among mixed isolates are considered in comparing outbreaks at different locations, the discriminatory power of the test increases

exponentially. Not only would the qualitative data work as a quick screening method to match unknown with reference samples, the quantitative data could greatly enhance the probative nature of the test results when comparing different isolates. In order to routinely utilize the quantitative data as an investigative tool however, the long term stability of the proportions of viral strains in a mixed sample would need to be investigated through serial passage of WSMV over several generations of outbreaks. Once tested, this two-pronged analysis method can easily be used for attribution of samples to an original party responsible for their cultivation or dissemination.

Of course, this research can only provide a theoretical hypothesis of the future of the assay's discriminatory power. Too few samples have been tested thus far and further validation would only be available after subjecting more isolates harvested from around the world to the assay.

Challenges of Forensic Testing with WSMV

Several important issues exist around testing WSMV for forensic purposes. One important factor to contemplate is WSMV's ubiquitous nature. Every year, wheat crops are affected by WSMV in a process that is independent of human-controlled methods. So the first, and potentially most important, issue associated with a WSMV outbreak is to distinguish between an intentional and a naturally occurring outbreak. Since natural outbreaks are common, the typical response involves simple cultural practices to preserve the growing and future crops in that particular field. The response elicited by a criminal act, however, would include the involvement of additional organizations and agencies that would never get involved during a natural outbreak.

Several characteristics could be taken into consideration to determine if an outbreak is one that had been intentionally set.

- Is the infection caused by an isolate known to be native to the area?
- Do other wheat growing farms in the vicinity also suffer from the outbreak?
- If other local farms are also inflicted with WSMV, is the entire outbreak caused by the same viral isolate?
- Is there a distinct increase in virulence or pathogenicity associated with this particular outbreak?
- Are there obvious signs of wheat curl mite invasion of the growing crop?
- Have the environmental conditions been favorable to the spread of the disease?
- Is the time of year conducive to a natural outbreak of WSMV?
- Does the outbreak follow a noticeable, non-wind borne pattern (like along the outside edge of the field, for instance)?
- Have any other recent environmental factors caused natural penetration points (wounds, etc.) for the virus?

While the preceding list is not meant to be exhaustive, any suspicions raised by the route of infection should require further investigation into its potential cause.

WSMV as a Model System

Inconspicuous infection with WSMV can be fairly difficult. Since the virus dies rapidly outside of a living host, has no natural means of active penetration of the plant, and must rely on wounding caused by environmental factors, mechanical abrasion

(through the use of agricultural machinery, e.g.), or the wheat curl mite to infect the plant, successfully infecting a field with WSMV would be a rigorous task. Such a full-scale outbreak would require artificially creating those wounds to allow the virus to even have a chance to penetrate its host. Successfully infecting an entire field of wheat would inevitably leave signs of human involvement in the outbreak. While the crops would still be infected, it would be difficult to keep the criminal act covert as human involvement would be discovered quickly, and the long-term incubation of WSMV that causes significantly more damage to the crop could be prevented. In terms of infections, far better and more self-sustaining pathogens exist that would require less elaborate means of contamination and have the ability to do far more damage to a greater range of crops.

However, that realization does not preclude the use of WSMV as a model system. Due to the long-standing history of WSMV as a known pathogen, hundreds of publications have gone into detail concerning numerous aspects of the disease. Knowledge of the pathogen's infectivity, reproduction rate, and method of transmission have been considered in the development of this assay, and the positive aspects of the research could readily be expanded outward to other similar pathogenic infections.

Sample Collection

The number of testable isolates is generally much higher than demonstrated through this research. One of the factors that hindered this study more than any other was the timing of certain research breakthroughs. Unfortunately, once a reliable testing system had been developed, the majority of naturally occurring outbreaks had already completed their yearly cycle. Shortly thereafter, finding reliable sources of new isolates

became increasingly difficult. The Forensic Sciences Department at the Center for Health Science of Oklahoma State University lacked the proper USDA clearance to import non-native WSMV isolates, which eliminated yet another route of obtaining additional isolates.

However, the potential for future screening of WSMV isolates based upon this method still exists. One recommendation for further investigation would propose that future researchers remain in contact with either a diagnostic laboratory or an extension agent that encounters natural outbreaks of WSMV on a yearly basis. With the first onset of WSMV infection, samples could be collected from around the state. Due to the small amount of green plant material required for the test, several samples could be collected during the peak infection time and then be properly stored for later genetic analysis. This approach could significantly increase the amount of isolates available for testing and more adequately assess the discriminatory power of the assay.

Potential Contaminants of the SNaPshot Reaction

Inspection of the potential mismatched binding sites from GenBank revealed a list of potential contaminants that only marginally matched any of the four primers used specifically for the SNaPshot™ multiplex assay. In fact, the influence of many of the contaminants (like *Mus musculus* (common house mouse), *Rattus norvegicus* (Norway rats), and *Pan troglodytes* (chimpanzee) for example) can essentially be negated as they have a minimal chance to contaminate the infected wheat before samples are taken.

Only one match provided greater than 58% complementarity to any given primer. That match—*Oryza sativa* at 88% to SNP-1—is a genomic DNA sequence from

chromosome three of a cultivar of domesticated rice originally cultivated in Japan. However, even at such a high match for the given primer, the rice mismatches five of the 45 nucleotides, making it an unlikely candidate to effectively compete for binding of the SNP-1 primer for WSMV. Through RNA and WSMV specific extraction, amplification, and purification protocols, however, something like genomic DNA of a Japanese rice cultivar should easily be eliminated as a potential threat to the veracity of the SNaPshot™ results.

CHAPTER V

CONCLUSION

Numerous adversities had to be overcome throughout the course of this research. However, these impediments provided more than just illogical results as they elucidated a truth universal to all experimentation: one should never blindly follow a protocol without questioning the science underlying every step. A skeptical approach to research can not only provide better understanding of methodologies but it can also prevent wasting time and money on procedures that may only serve to confuse the results. Nevertheless, this approach did effectively negate several potential assays before eventually leading to the SNaPshot™ kit.

The SNP-specific assay developed through the course of this research has the ability to quell one facet of an ever-growing threat of agroterrorism. Three isolates of different origin were examined and the assay discriminated all of them both qualitatively and quantitatively. If properly expanded, this assay would provide investigators with an invaluable tool to decisively establish a link between a viral sample seized from a clandestine laboratory and a viral sample collected from an intentionally administered field outbreak. Likewise, the opportunity still exists to even expand upon this research by identifying and synthesizing additional SNP primers to add to the multiplexed SNaPshot™ reaction.

In summary, the answers to the proposed research questions are:

1. Comparing data obtained by sequencing the coat protein region of the three isolates against the Sidney-81 consensus sequence shows that distinct SNPs do exist and differentiate the various WSMV isolates.
2. An assay was developed that relied upon RT-PCR, Agarose Gel Electrophoresis, PCR, and Genetic Analysis which can all be performed in a modern DNA analysis laboratory.
3. The developed assay not only effectively discriminated between isolates qualitatively as expected, but it also had a component that quantitatively discriminated differing isolates as well. The entire process can be completed in less than 48 hours from the start of extraction.
4. An additional primer that served as a diagnostic, internal control was synthesized and added to the multiplexed SNaPshot™ reaction to ensure that the results were indeed amplified from WSMV.

REFERENCES

Applied Biosystems, Inc. (2000). ABI Prism© SNaPshot™ Multiplex Kit. *Protocols: Applied Biosystems, Inc.*, 1.

Burges, Nichole, Jones, R., & Kumar, S. (2003). Wheat Streak Mosaic Virus. *Factsheet: Government of Western Australia, 5,* 1-2.

Chenault, Kelly D., Hunger, R.M., & Sherwood, J.L. (1996). Comparison of the Nucleotide Sequence of the Coat Protein Open Reading Frame of Nine Isolates of Wheat Streak Mosaic Rymovirus. *Virus Genes, 13(3),* 187-188.

Decker-Walters, Deena S., Chung, S-M., & Staub, J.E. (2004). Plastid Sequence Evolution: A New Pattern of Nucleotide Substitution in the Cucurbitaceae. *Journal of Molecular Evolution, 58(5)*, 606-614.

French, Roy, & Stenger, D.C. (2003). Evolution of Wheat Streak Mosaic Virus: Dynamics of Population Growth Within Plants May Explain Limited Variation. *Annual Review of Phytopathology, 41,* 199-214.

French, Roy, & Stenger, D.C. (2005). Population structure within lineages of Wheat streak mosaic virus derived from a common founding event exhibits stochastic variation inconsistent with the deterministic quasi-species model. *Virology, 343,* 179-189.

Gould, Ernest A. (1999). Methods for Long-Term Virus Preservation. *Molecular Biotechnology, 13,* 57-66.

Hall, Jeffrey S., French, R., Morris, T.J., & Stenger, D.C. (2001). Structure and Temporal Dynamics of Populations within Wheat Streak Mosaic Virus Isolates. *Journal of Virology, 75(21)*, 10231-10243.

Hammon, B. (2003). Management of High Plains Disease in Western Colorado Sweet Corn. *Extension Publication: Colorado State University*, 1.

Hunger, R.M., Sherwood, J.L., Evans, C.K., & Montana, J.R. (1992). Effects of Planting Date and Inoculation Date on Severity of Wheat Streak Mosaic in Hard Red Winter Wheat Cultivars. *Plant Dis., 76,* 1056-1060.

Lab Rat, The (2005). HpaII Restriction Enzyme. Retrieved September 2, 2007 from The Lab Rat Web site: http://www.thelabrat.com/restriction/HpaII.shtml

Lab Rat, The (2005). AvaI Restriction Enzyme. Retrieved September 2, 2007 from The Lab Rat Web site: http://www.thelabrat.com/restriction/AvaI.shtml

McNeil, Jill E., French, R., Hein, G.L., Baenziger, P.S., & Eskridge, K.M. (1996). Characterization of Genetic Variability Among Natural Populations of Wheat Streak Mosaic Virus. *Phytopathology*, *86*, 1222-1227.

Monie, Daphne, Simmons, R.P., Nettles, R.E., et al. (2005). A Novel Assay Allows Genotyping of the Latent Reservoir for Human Immunodeficiency Virus Type 1 in the Resting CD4+ T Cells of Viremic Patients. *Journal of Virology*, 79(8), 5185-5202.

ICTVdB Management (2006). Wheat Streak Mosaic Virus. *ICTVdB – The Universal Virus Database,* version 4. 25 April 2006. 2 September 2007. <http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/00.057.0.06.001.htm>

Montana, J.R., Hunger, R.M., & Sherwood, J.L. (1996). Serological Characterization of Wheat Sreak Msoaic Virus Isolates. *Plant Dis., 80,* 1239-1244.

Murray, Gordon M., Knihinicki, D., Wratten, K., & Edwards, J. (2005). Wheat streak mosaic and the wheat curl mite. *Primefacts: NSW Department of Primary Industries, 99,* 1-4.

Pear, Robert (2004). U.S. Health Chief, Stepping Down, Issues Warning. *New York Times, 24 Dec. 2004, natl. ed.*

Rabenstein, Frank, Seifers, D.L., Schubert, J., French, R., & Stenger, D.C. (2002). Phylogenetic relationships, strain diversity and biogeography of tritimoviruses. *Journal of General Virology, 83,* 895-906.

Rothamsted Research (1994). Electron Micrographs of Plant Viruses. Retrieved September 2, 2007 from the homepage of Rothamsted Research: http://www.rothamsted.ac.uk/ppi/links/pplinks/virusems/a1.gif

Sherwood, J.L., Hunger, R.M., Keyser, G.C., & Myers, L.D. (1990). Production of a Monoclonal Antibody for Evaluation of Hard Red Winter Wheat Cultivars to Wheat Streak Mosaic Virus. *Food & Agricultural Immunology, 2,* 155-161.

Sill, Jr., W.H. (1953). Some Characteristics of the Wheat Streak-mosaic Virus and Disease. *Transactions of the Kansas Academy of Science, 56(4),* 418-424.

Sloderbeck, Phillip E. (1995). Wheat Curl Mite. *Entomology Wheat Insects: Kansas State University*, 1.

Somsen, H.W., & Sill, Jr., W.H. (1970). The Wheat Curl Mite, *Aceria tulipae* Keifer, in Relation to Epidemiology and Control of Wheat Streak Mosaic. *Research Publications: Kansas State University – Agricultural Experiment Station*, *162,* 1-24.

Stack, J., Cardwell, K, Hammerschmidt, R., Byrne, J., et al. (2006). The National Plant Diagnostic Network. *Plant Disease, 90(2),* 128-136.

Stenger, Drake C., Seifers, D.L., & French, R. (2002). Patterns of Polymorphism in *Wheat streak mosaic virus*: Sequence Space Explored by a Clade of Closely Related Viral Genotypes Rivals That between the Most Divergent Strains. *Virology, 302,* 58-70.

Stenger, Drake C., Hein, G.L., Gildow, F.E., Horken, K.M., & French, R. (2005). Plant Virus HC-Pro Is a Determinant of Eriophyid Mite Transmission. *Journal of Virology, 79(14),* 9054-9061.

Townsend, Lee, Johnson, D., & Hershman, D. (1996). Wheat Streak Mosaic Virus and the Wheat Curl Mite. *ENTFACTS: University of Kentucky,* 1-3.

United States Department of Agriculture (2007). *Data Set: State Fact Sheet (United States).* Retrieved September 2, 2007 from the United States Department of Agriculture Web site: http://www.ers.usda.gov/StateFacts/US.HTM

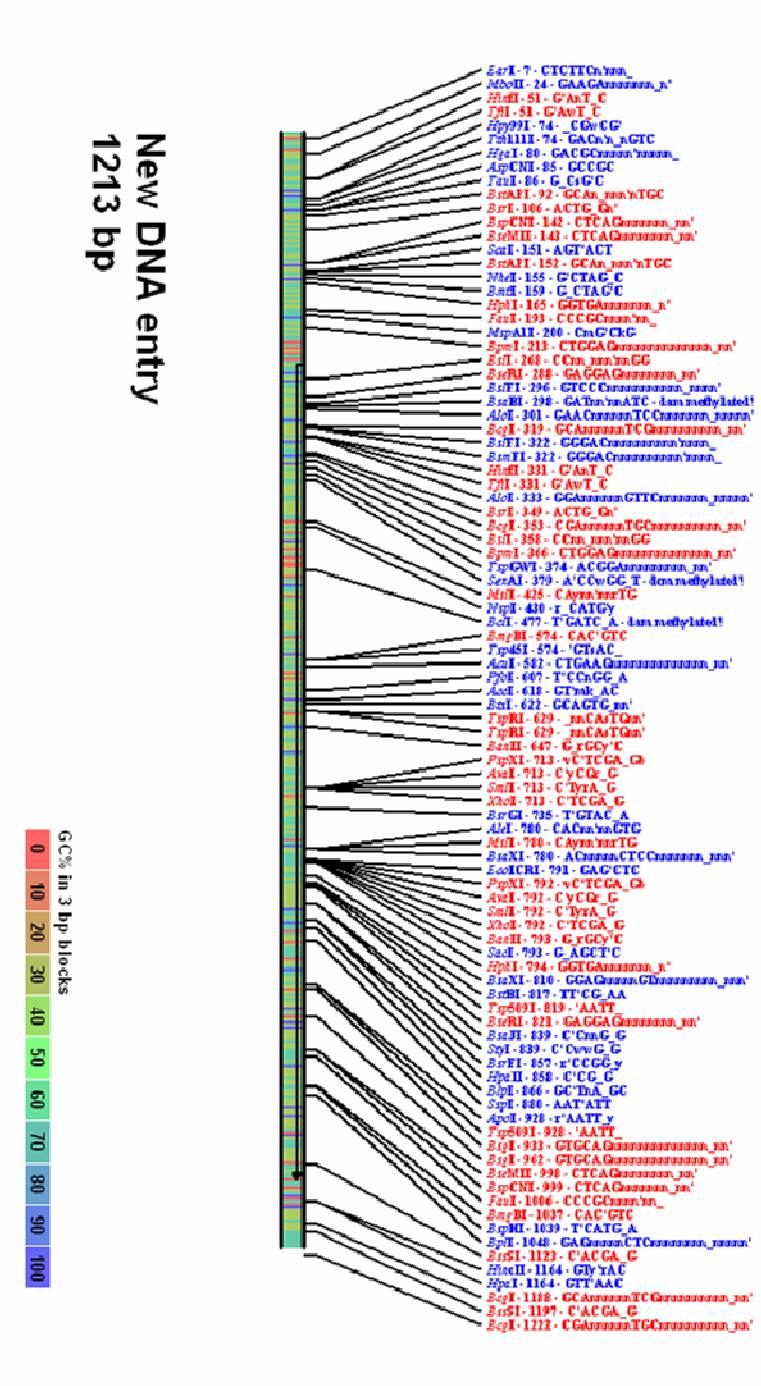
Venter, J. Craig, Adams, M.D., Myers, E.W., Li, P.W., et al. (2001). The Sequence of the Human Genome. *Science, 291(5507),* 1304-1351.

Watkins, John E. (2002). Wheat Disease Fact Sheet No. 1-7 Management Programs for Wheat. *University of Nebraska Publications: NebFact*, 1-7.

Ziebolz, B. & Droege, M. (2007). Toward a new era in sequencing. *Biotechnology Annual Review, 13,* 1-26.

APPENDIX

APPENDIX A – Restriction Map of WSMV



APPENDIX B – Single Nucleotide Sequencing Electropherograms

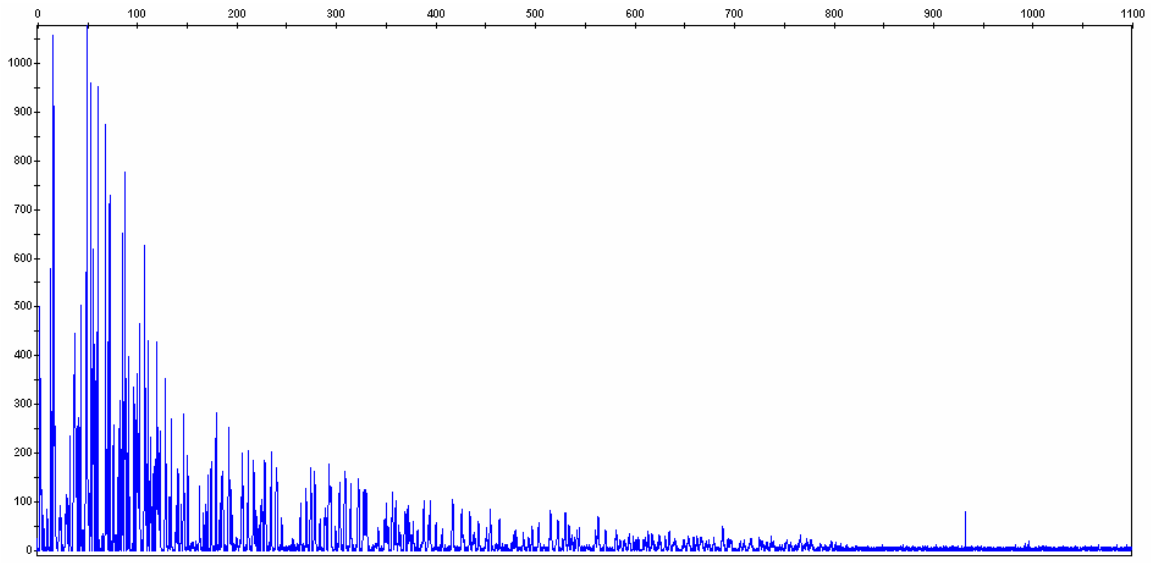


Figure B-1. Plant C with Labeled ddGTP and Unlabeled Primer Electropherogram.

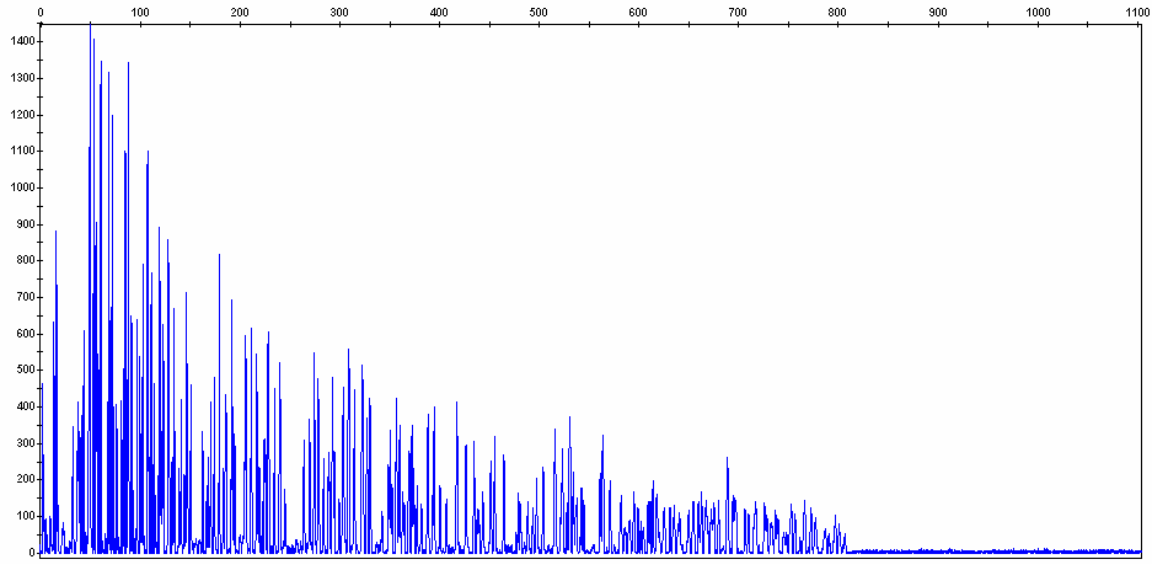


Figure B-2. Isolate 964 with Labeled ddGTP and Unlabeled Primer Electropherogram.

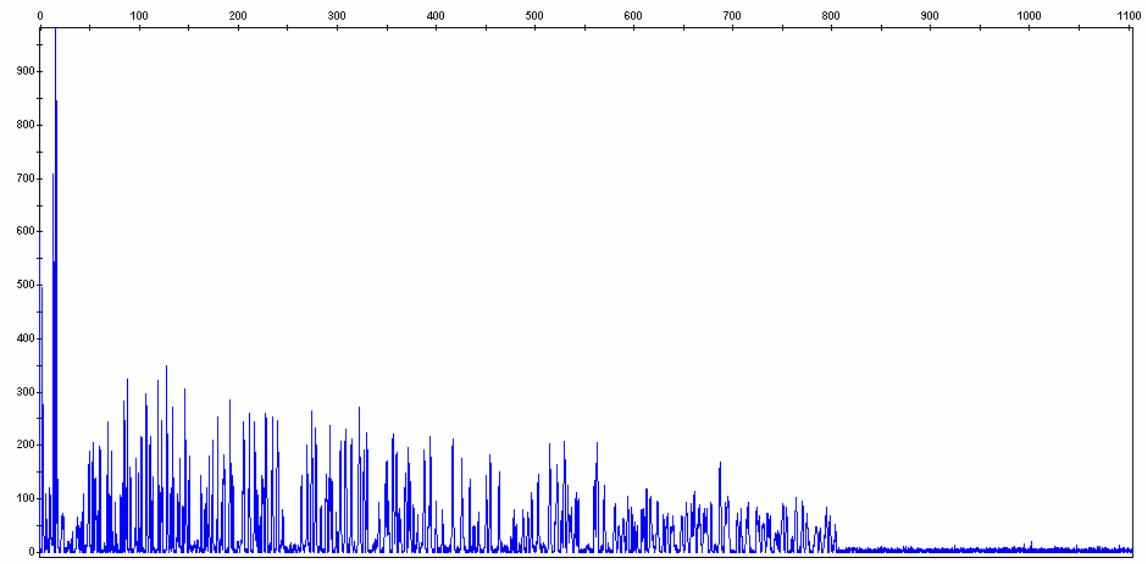


Figure B-3. OSU with Labeled ddGTP and Unlabeled Primer Electropherogram.

APPENDIX C – Complete Nucleotide Sequencing Electropherograms

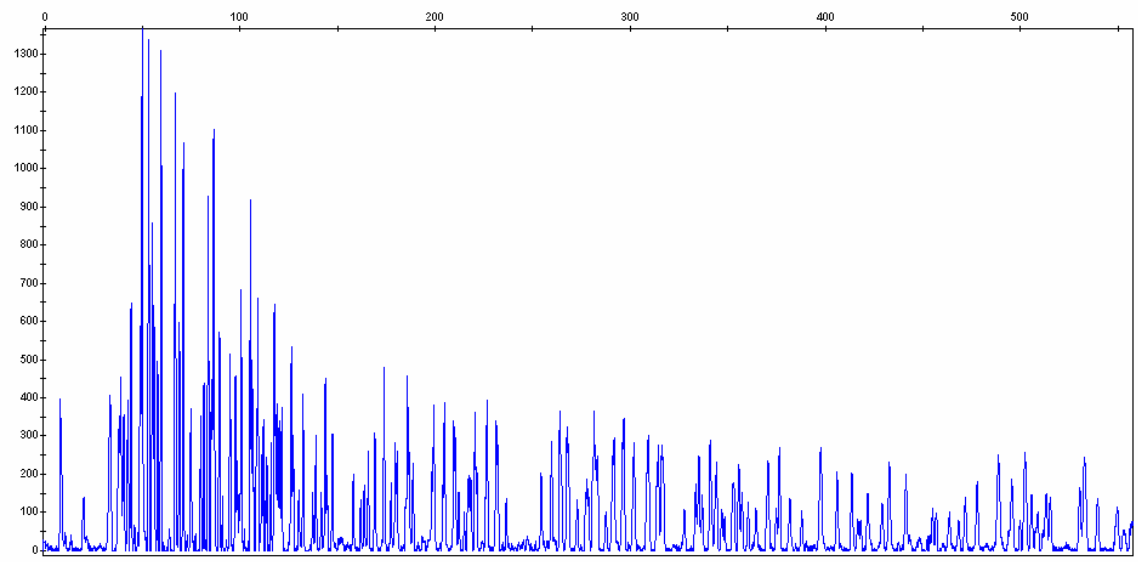


Figure C-1. Plant C Electropherogram Results from BigDye® Terminator Kit.

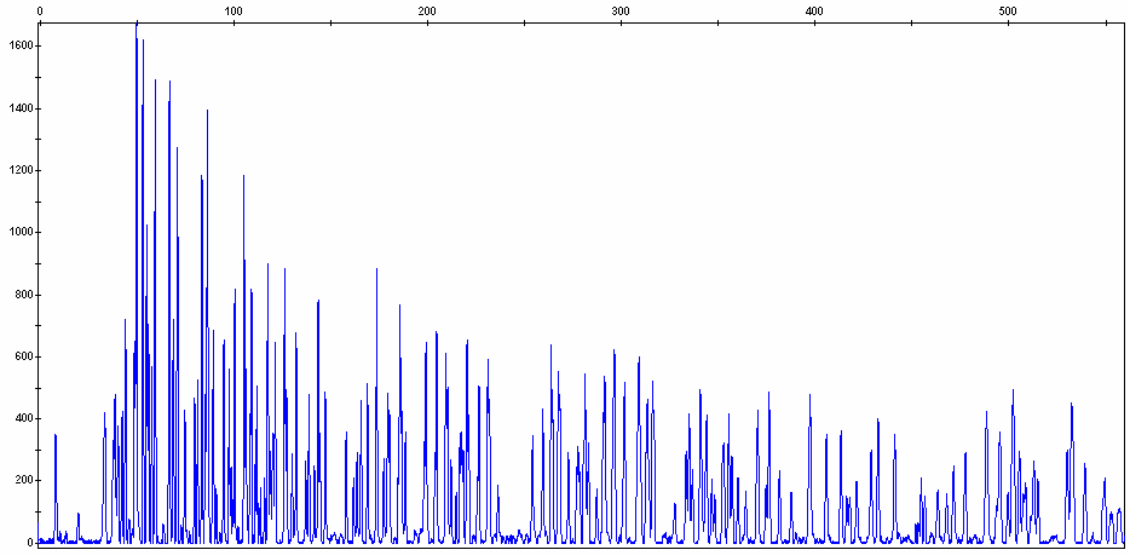


Figure C-2. Isolate 964 Electropherogram Results from BigDye® Terminator Kit.

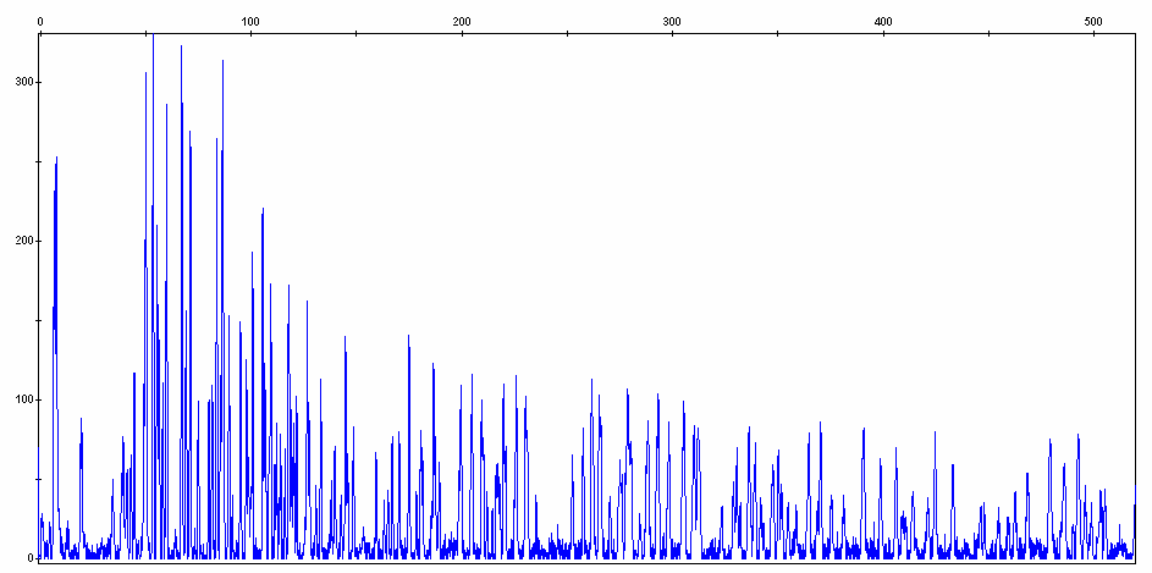


Figure C-3. OSU Electropherogram Results from BigDye® Terminator Kit.

APPENDIX D – Repeated Runs of Multiplexed SNaPshot Reactions

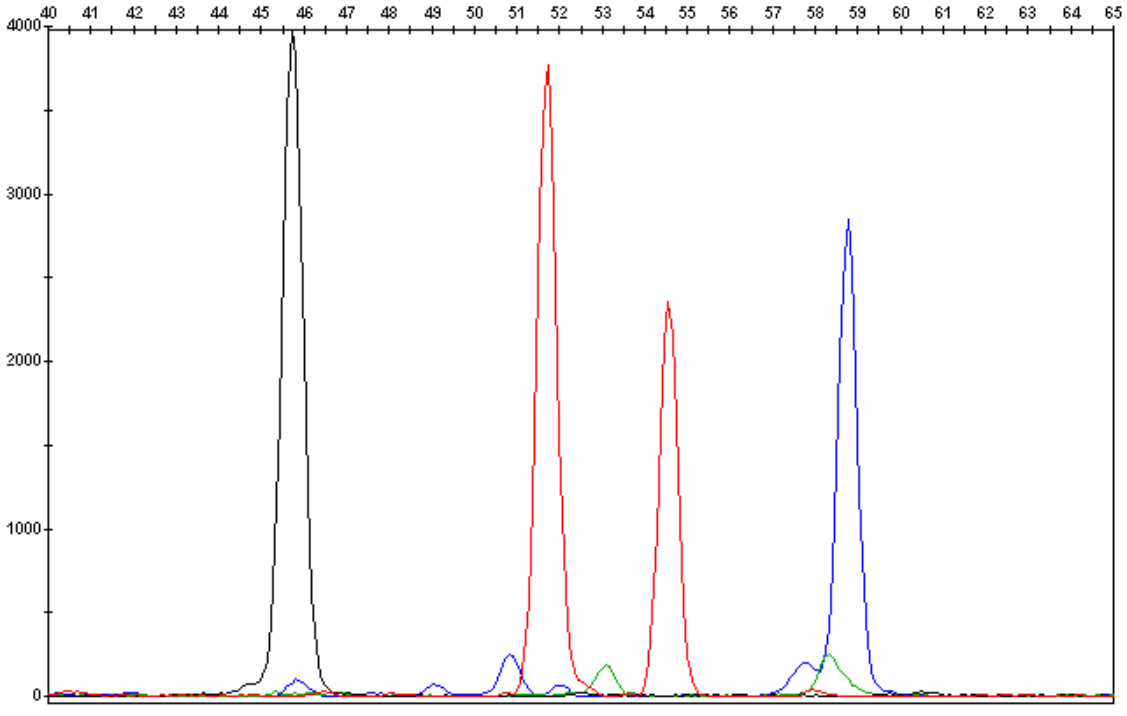


Figure D-1. SNaPshot™ Electropherogram of Plant C, First Extraction.

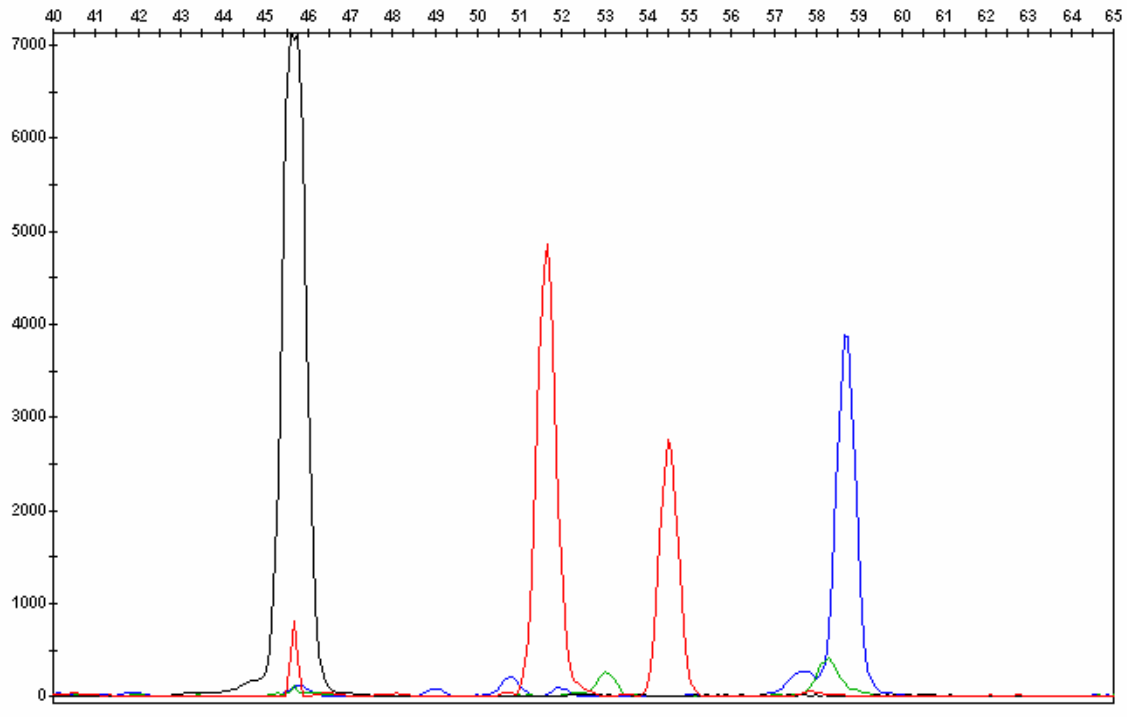


Figure D-2. SNaPshot™ Electropherogram of Plant C, Second Extraction.

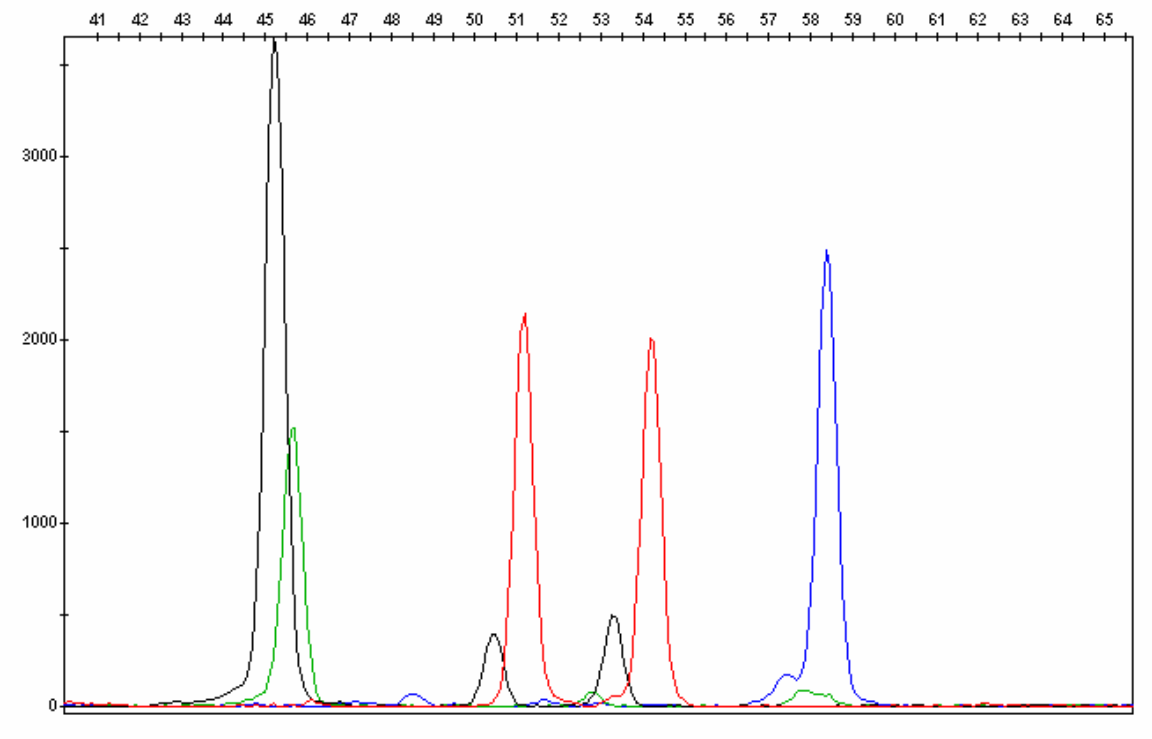


Figure D-3. SNaPshot™ Electropherogram of Isolate 964, First Extraction.

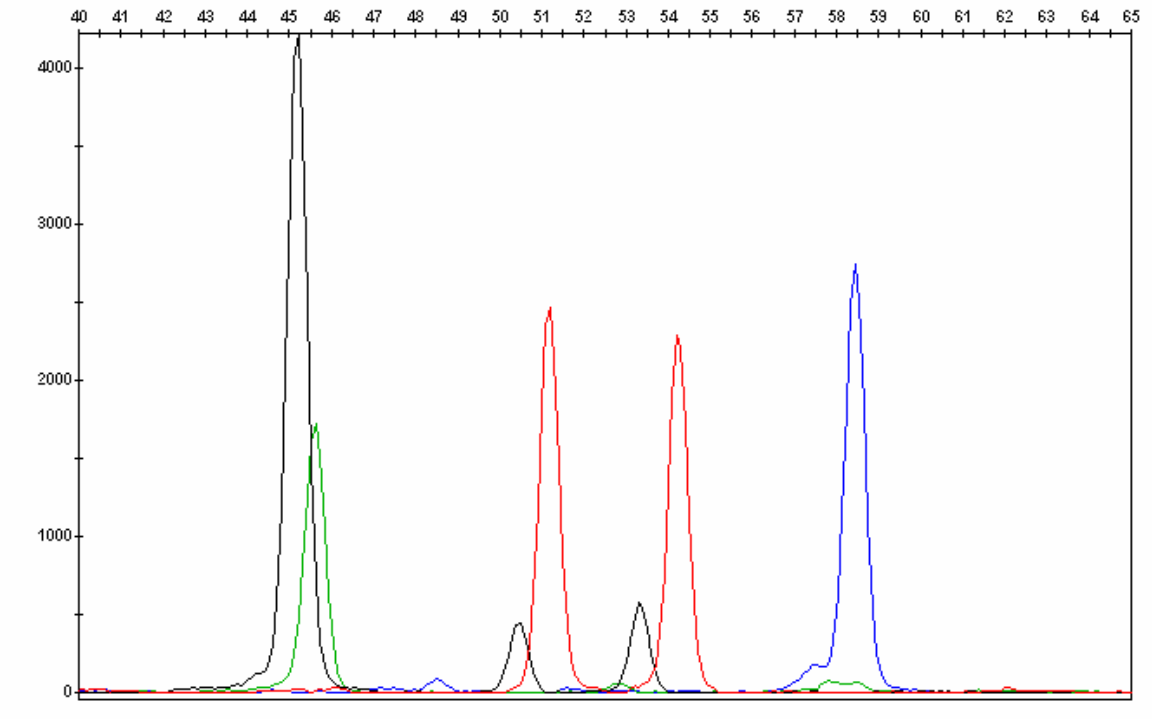


Figure D-4. SNaPshot™ Electropherogram of Isolate 964, Second Extraction.

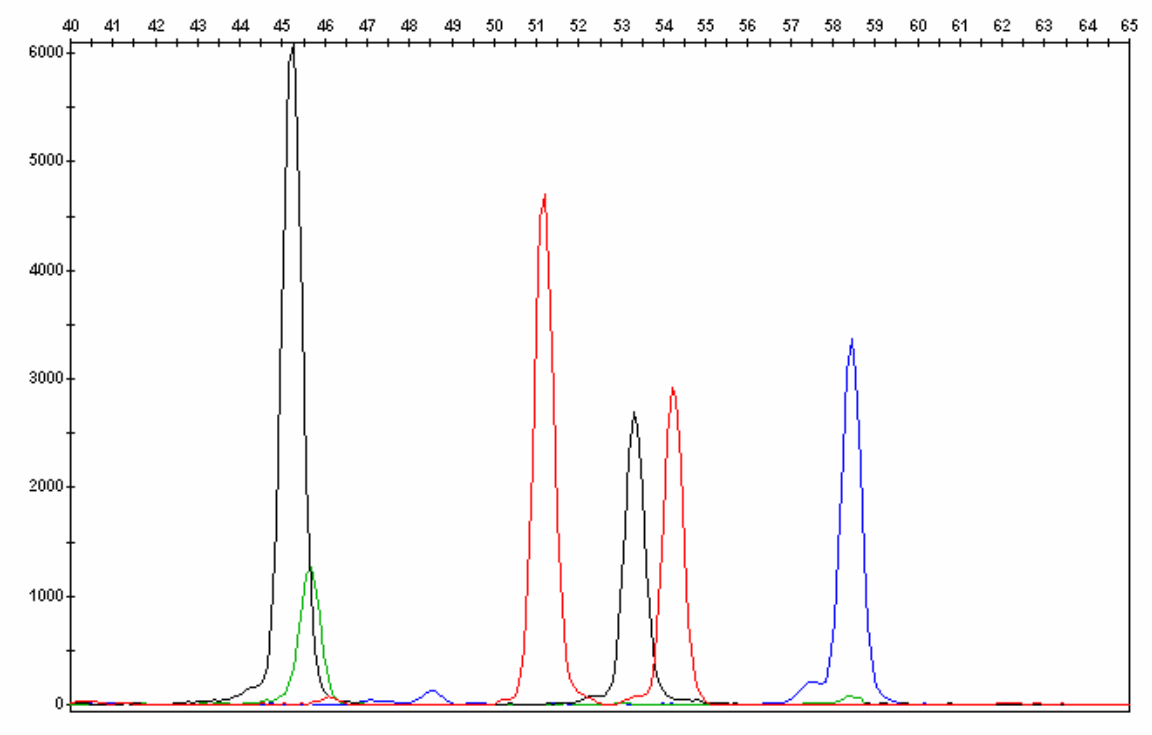


Figure D-5. SNaPshot™ Electropherogram of OSU, First Extraction.

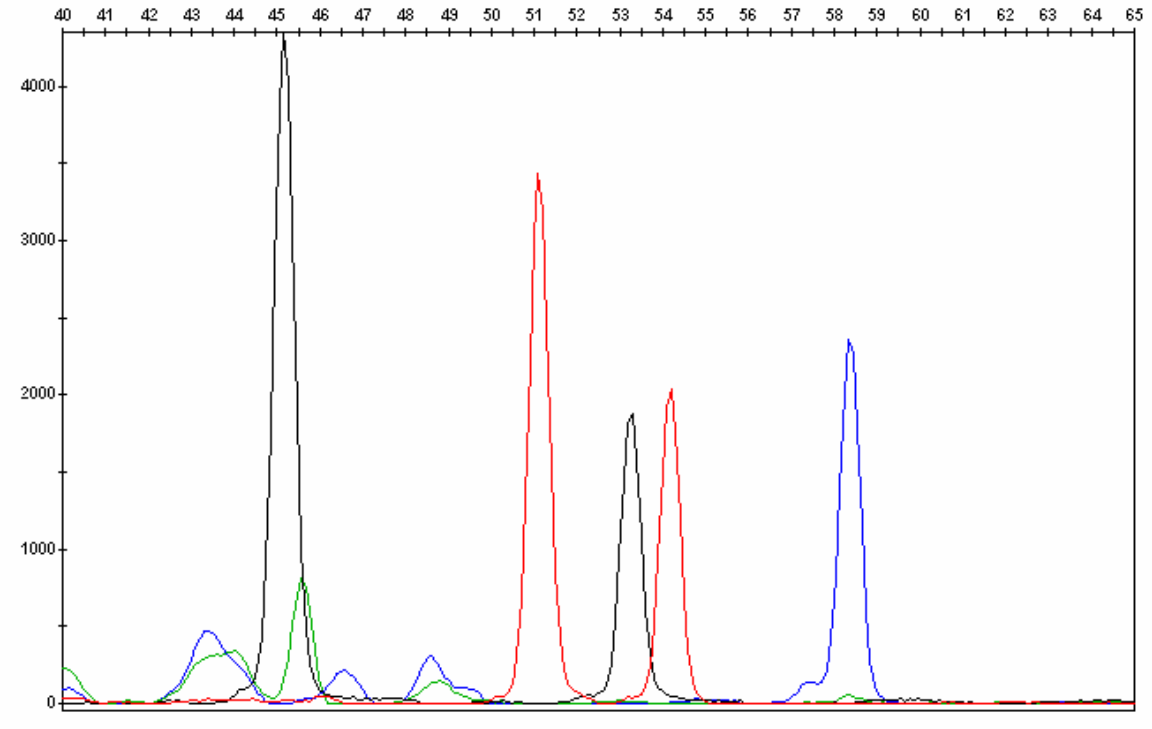


Figure D-6. SNaPshot™ Electropherogram of OSU, Second Extraction.

APPENDIX E – Negative Control Data Electropherograms

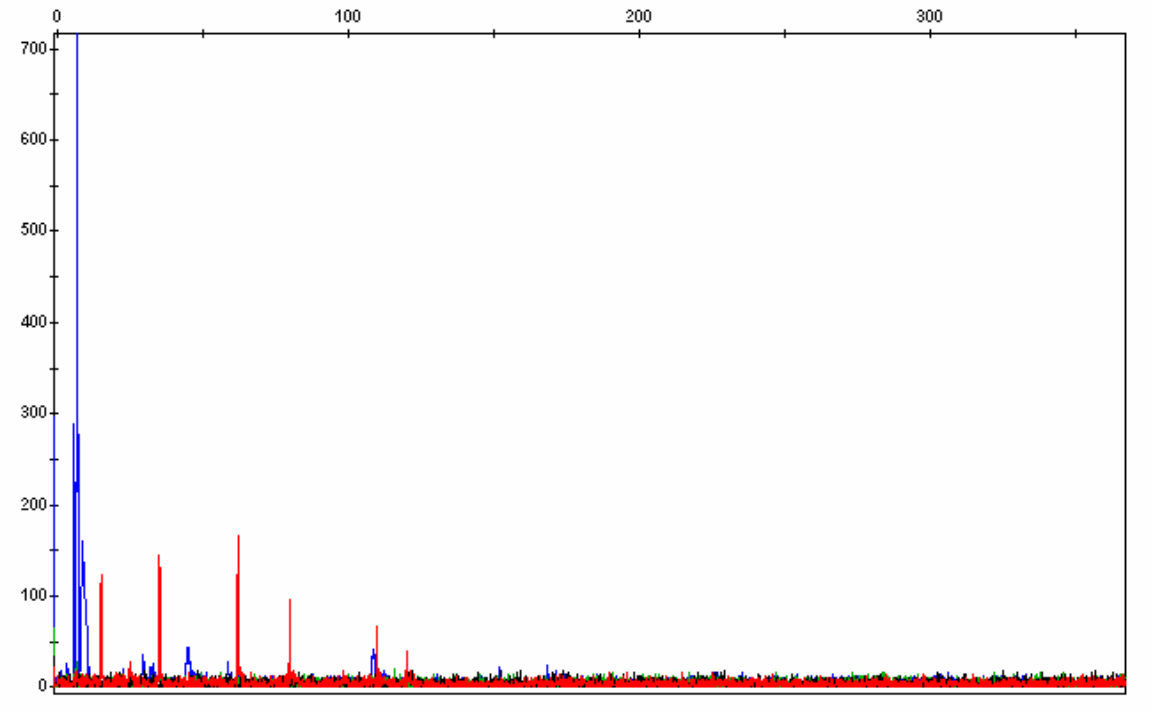


Figure E-1. SNaPshot™ Electropherogram of BSMV, Alpha Isolate

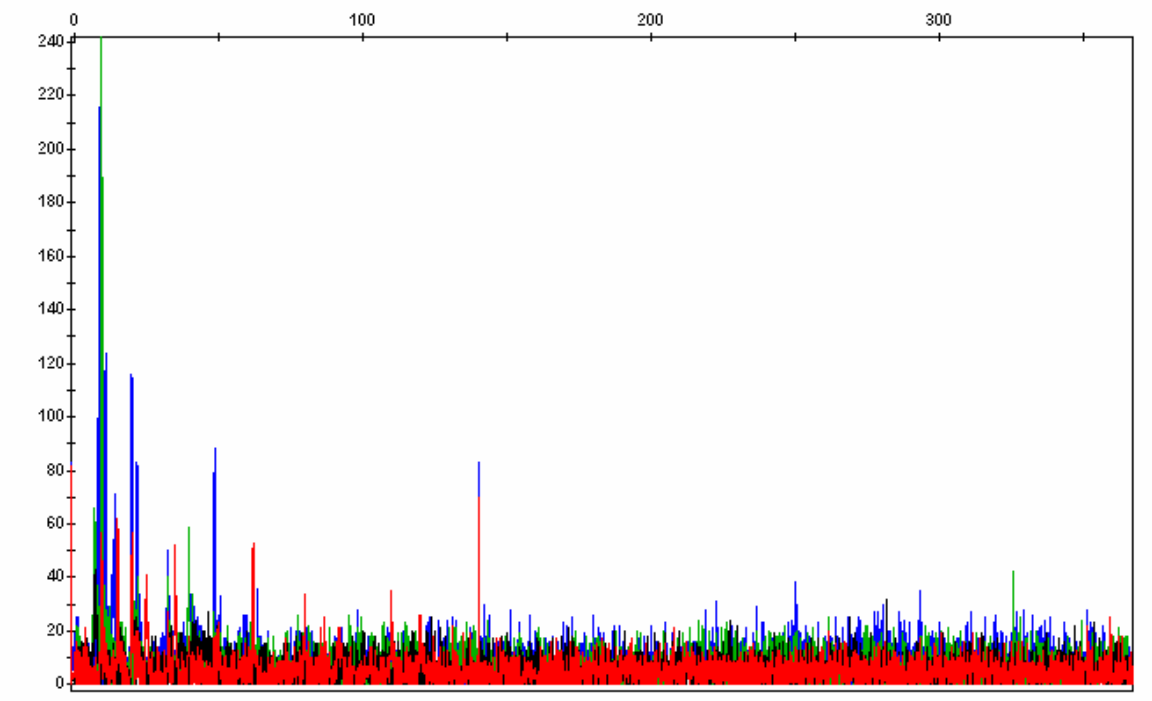


Figure E-2. SNaPshot™ Electropherogram of BSMV, Beta Isolate

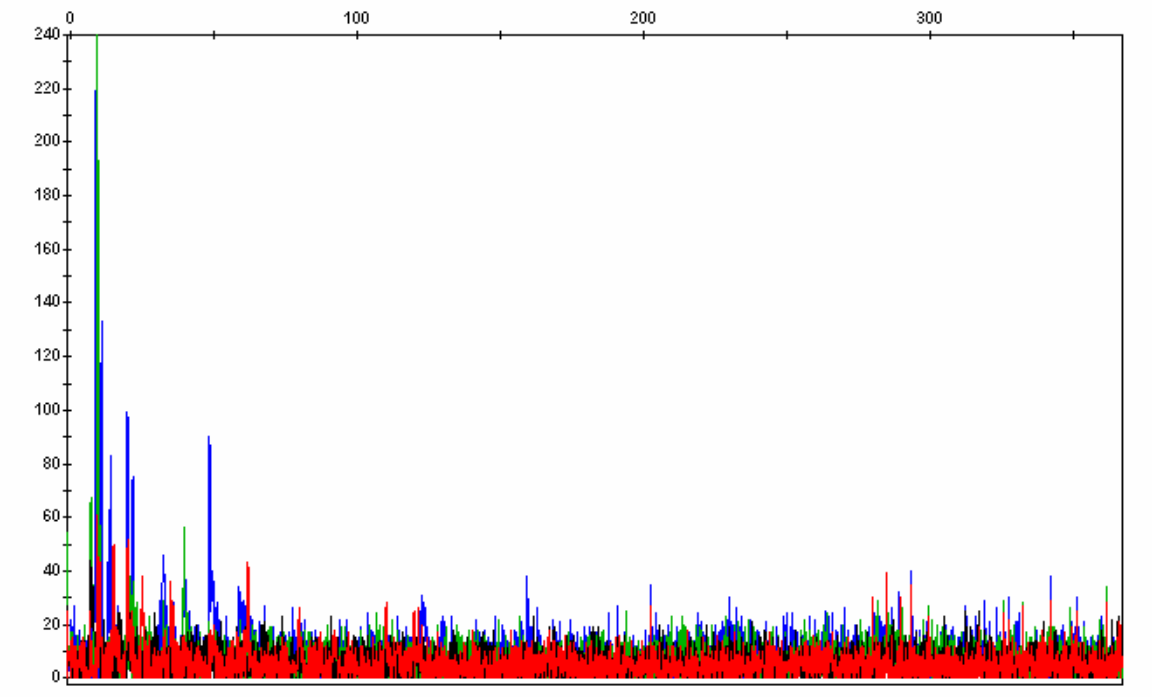


Figure E-3. SNaPshot™ Electropherogram of BSMC, Gamma Isolate

VITA

Jesse David Carver

Candidate for the Degree of

Master of Science

Thesis: MOLECULAR TYPING OF WHEAT STREAK MOSAIC VIRUS FOR FORENSIC APPLICATIONS

Major Field: Forensic Science

Biographical:

Personal Data: Born in Downey, California, on April 5, 1983, the son of Bura and Lois Carver.

Education: Graduated from Southside High School in Fort Smith, Arkansas in May, 2001.

Received a Bachelor of Science degree with a major in Biological Sciences and minors in Anthropology and Mathematics from the University of Tulsa, Tulsa, Oklahoma in May, 2005.

Completed the requirements for the Master of Science in Forensic Science at Oklahoma State University Center for Health Sciences with a concentration in Forensic Identification with DNA in December, 2007.

Name: Jesse David Carver Date of Degree: December, 2007

Institution: Oklahoma State University-CHS Location: Tulsa, Oklahoma

Title of Study: MOLECULAR TYPING OF WHEAT STREAK MOSAIC VIRUS FOR FORENSIC APPLICATIONS

Pages in Study: 85 Candidate for the Degree of Master of Science

Major Field: Forensic Identification with DNA

Scope and Method of Study: The field of agricultural biosecurity combines traditional aspects of forensic science with plant pathology. With the threat of agroterrorism growing, steps must be taken in many areas to establish preparedness and to train enforcement personnel and extension agents in tracking and prosecuting the responsible individuals. The purpose of this study was to use Wheat Streak Mosaic Virus as a model system to develop a molecular assay using current, popular forensic laboratory equipment that can effectively compare samples collected from a suspected agroterrorism event against reference samples collected from a potentially responsible clandestine laboratory. Viral RNA extractions were performed with the MagMAX™ Kit (Ambion, Inc., Foster City, CA). cDNA synthesis and viral genome amplification were performed with the SuperScript™ One-Step RT-PCR Kit (Invitrogen, Inc., Carlsbad, CA). Single nucleotide polymorphisms were identified from sequencing data of three WSMV isolates. Three primers based upon these SNPs and a fourth primer included as an internal control and diagnostic marker were synthesized and the SNaPshot™ Kit (ABI, Inc., Foster City, CA) was used to discriminate the three WSMV isolates.

Findings and Conclusions: The three SNP-specific primers used during SNaPshot™ analysis showed distinct qualitative differences between the three WSMV isolates tested. The fourth, internal control primer produced a positive result in every test, confirming the presence of WSMV cDNA within each sample. However, additional, unexpected peaks occurred at various sites in the electropherograms. After repeated SNaPshot™ reproducibility tests, blank assays, and negative controls, these additional peaks were determined to be evidence of mixed populations of WSMV within two of the three infections. The additional peaks therefore provided a second test for attribution that relied on the quantitative level of infection with genetically distinct isolates of WSMV instead of solely relying on the qualitative polymorphisms associated with an infection. This increased the potential discriminatory power of the assay exponentially. Future study with this assay would allow further testing of the reproducibility of the current isolates, subject previously uninvestigated natural isolates from around the world, and incorporate additional SNP-specific primers.

ADVISER'S APPROVAL: Dr. Robert W. Allen