**SZENT ISTVÁN UNIVERSITY**
**FACULTY OF AGRICULTURAL AND ENVIRONMENTAL SCIENCES**
**MSC IN AGRICULTURAL BIOTECHNOLOGY**

# Investigation of GRVFV variants in grapevine at a plantation near Pécs

**Author:**
Dino Muratović

**Primary Thesis Adviser:**
Dr. Erzsebet Kiss
**Institute/Department**
Institute of Genetics and Biotechnology

**Independent Consultant:**
Dr. Éva Várallyay

Gödöllő
2018

**Contents**

## 1. INTRODUCTION

Grapevine (*Vitis spp.*) is one of the most widely grown fruit crops worldwide. Recent worldwide production estimates (2014) are 7.6 million hectares and 74 milion metric tons. Grapevine finds it's main uses for wine production, but also for fresh fruit, raisins, juice, vinegar, seed oils and several other products.

Grapevines belong to the family Vitaceae, which contains 12 genera and over 700 species. Most species within this family are climbing vines and include genera such as *Ampelocissus, Ampelopsis, Cayratia, Cissus, Clematicissus, Parthenocissus, Tetrastigma* and *Vitis*.

Since these species find use in many areas of industrial production, importance of early an detection of any type of virus is important for keeping the grapevine plantations virus free.

Thanks to the attention that has been given to virus detection in many plants, including grapevines, development of new techniques has given a rise to possibility of fast, sensitive and accurate detection of various viruses.

This research was based on investigation of *Grapevine rupestris vein feathering virus* (GRVFV) that was present at a plantation near Pecs.

# 2. LITERATURE REVIEW

## 2.1 Evolution of Vitis species

A fairly large number of *Vitis* species (approximately 60) have evolved worldwide, of which *V. vinifera* has become the most widespread for wine and table use (Meng et al., 2017).

Europe and Central Asia has a single species, *V. vinifera*, which is subdivided into *V. vinifera ssp. sativa* (cultivated grape) and *V.vinifera ssp. Sylvestris typica* (wild grape) (Olmo, 1996). Cultivars that are present nowadays likely arose initially from collecting and planting seeds. These seeds propagated populations would have been highly heterozygous, and some specific cultivars could not have been selected until vegetative propagation has been done (by cuttings or layerings) (Meng et al., 2017).

Cultivated grapevines (*Vitis vinifera spp. sativa*) are thought to have been domesticated from wild populations of *Vitis vinifera spp. sylvestris* (Meng et al., 2017). These wild vines are dioecious plants that are still occurring in small isolated populations along riverbank forests from the Atlantic coast of Europe to Tajikistan and western Himalayas (Zohary and Hopf, 2000).

## 2.2 Grapevine viruses

Around 70 viruses and virus-like diseases are reported (Martelli, 2014) from grapevines, which are characterized by a wide variety of symptoms: malformations of leaves and twigs, foliar discolorations (reddening, yellowing, chlorotic or bright yellow mottling, ringspots and also line patterns), grooving and/or pitting of the woody cylinder, delayed bud break, stunting and decline (Meng et al., 2017). All of this can influence the productive lifespan of the vineyards, which can be shortened, and the quantity and quality of the crop is also badly affected.

Prevailing agents of the three major disease complexes (infectious degeneration/decline, leafroll and rugose wood) are either viruses with isometric particles, the most relevant of which are transmitted by nematodes (nepoviruses) or viruses with filamentous particles, which are transmitted by pseudococcid mealybugs and soft scale insects (closteroviruses and vitivurses). So far, no vectors are known for the viruses of a fourth complex (fleck).

## 2.3 Infectious Degeneration/Decline

Recognized as putative agents of infectious defenration/decline are viruses with isometric particles classified in the genus *Nepovirus* (Figure 1), (exception is Strawberry latent ringspot virus, which is unassigned member of the family *Secoviridae*), many of which have a recognized nematode vector (Meng et al., 2017). These viruses have bipartite, single-stranded, positive-sense RNA genomes. The complete sequence of 12 of them has been previously determined (Martelli, 2014).

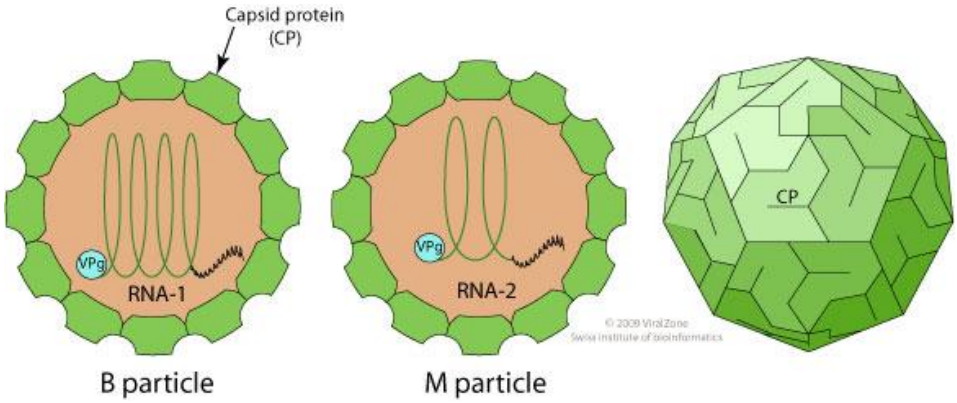



**Figure 1. Structure of Nepovirus virion (Taken from: https://viralzone.expasy.org/300?outline=all_by_species)**

Viruses that are involved in degenerative diseases (fanleaf and like) are reffered to as Old World nepoviruses, with the exception of GFLV, because they occur in this geographical area and have vectors sharing the same territorial distribution (Martelli and Taylor, 1990). But, degenerative diseases and relative agents are prevailing in Continental and Mediterranean Europe and it is believed that these areas are most likely their place of origin, whilst other diseases denoted „grapevine decline“, the eliciting viruses and vectors are found primarily in North America (Meng et al., 2017).

## 2.4 Leafroll

Grapevine leafroll disease has been described throughout different regions of Europe and elsewhere for over a century (Hoefert and Gifford, 1967) and was first shown to be transmissible to vines in 1936 (Scheu, 1936).

Grapevine leafroll disease (GLD) is most obvious and problematic in cool-climate regions, where fruit of infected vines has delayed ripening that further results in lowered brix which affects wine quality (Over de Linden and Chamberlain, 1970). GLD symptoms usually appear in the fall, when red cultivars display leaf reddening with green venation. These symptoms might not be as apparanet in white cultivars, but there is a slight leaf chlorosis. Both, red and white, cultivars develop downward rolling of the leaf margins and phloem is disrupted. More significant losses result from combination of factors, which can include yield reductions of up to 40% increased management costs, shortened vineyard life spans, and also impacts on wine quality which are direct result of decreased fruit quality and delayed maturation (Woodrum et al., 1984).

GLD has three essential biological components:

1. A complex of viruses in the *Closteroviridae*
2. Grapevine host plants
3. Species of mealybugs (*Pseudococciade*) and soft scales (*Coccidae*)

Virus species causing GLD are named Grapevine leafroll-associated virus 1, Grapevine leafroll-associated virus 2 and so on. All of the GLRaVs belong to the genus of *Ampelovirus*, with exception of GLRaV-a and GLRaV-7. GLRaVs in the *Ampelovirus* genus are divided into two phylogenetic groups, one of which includes GLRaV-4, -5, -6, -9 and another comprising GLRaV-1 and -3 (Maliogka et al., 2009).

Both groups of GLRaV ampeloviruses, are filamentous virions (Figure 2) with a large (13-18 kb) positive-sense single-stranded RNA genome (Figure 3), (Fuchs et al., 2009).

Some differences between genomes of the two groups have been noticed. The genomes of GLRaV-4-like species are around 5 kb smaller and lack several open reading frames on their 3' ends that are present in GLRaV-1 and -3 (Thompson et al., 2012).


**Figure 2. Virion of Ampelovirus**

**(Taken from: https://viralzone.expasy.org/285?outline=all_by_species)**

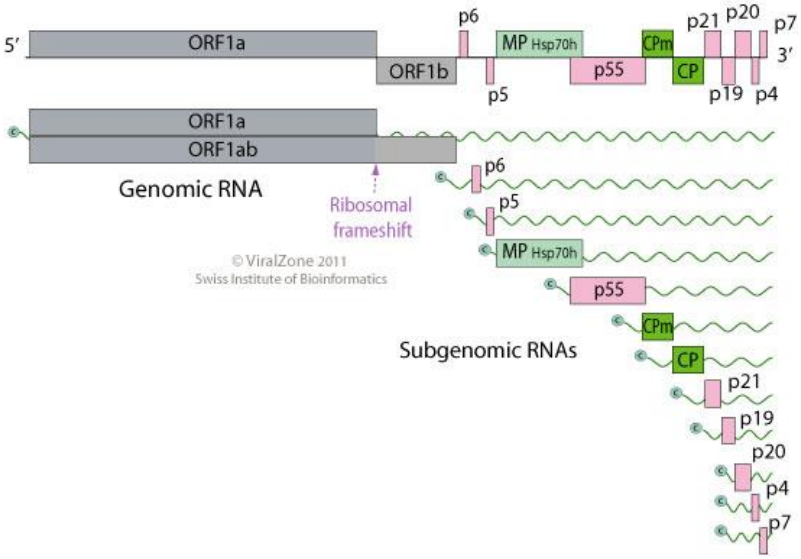



**Figure 3. Ampelovirus genome**

**(Taken from: https://viralzone.expasy.org/285?outline=all_by_species)**

## 2.5 Rugose wood

Rugose wood is a graft-transmissible disease first reported from Italy and soon afterward from Hungary (Martelli et al., 1967), and it's a complex disorder whithin wich, based on the differential reactions of the indicators *V. rupestris*, LN33, and Kober 5BB (Savino et al., 1989), four different syndromes have been identified:

1. Rupestris stem pitting (RSP)
2. Kober stem grooving (KSG)
3. Corky bark (CB)
4. LN-33 stem grooving (LNSG)

Rugose wood genome is consisted of a linear ssRNA(+), where 3' terminus is polyadenylated while 5' is capped (Figure 4). A breakthrough in rugose wood epidemiology came when GVA was, through experiment, transmitted by *Pseudococcus longispinus* (Rosciglione et al., 1983). This was the first evidence that pseudococcis mealybugs, till then known only as DNA virus vectors, were able to transmit also RNA viruses.

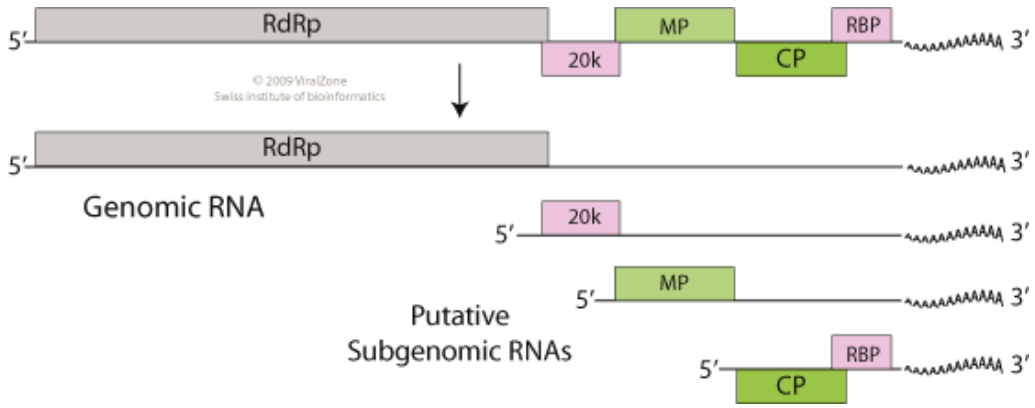



**Figure 4. Vitivirus genome**

**(Taken from: https://viralzone.expasy.org/270?outline=all_by_species)**

## 2.6 Fleck

Fleck, which represents a disease with distrubution all over the world, is latent in European grape cultivars and in most American rootstocks (Meng et al., 2017). Characteristic symptoms that are expressed in *V. rupestris* consist of clearing of the veins of third and fourth order resulting in localized translucent spots. Leaves with intense flecking are wrinkled, twisted and can curl upward (Hewwit et al., 1962, 1972).

Common agent is *Grapevine fleck virus* (GFkV), species that belong to genus *Maculavirus* (Martelli et al., 2014). It has isometric particles with rounded contour and a surface structure that is prominent and contains a single-stranded, positive-sense RNA genome (Figure 5), (Meng et al., 2017). These properties are also shared by three additional viruses: *Grapevine asteroid mosaic-associated virus* (GAMaV), *Grapevine rupestris vein feathering virus* (GRVFV) and *Grapevine redglobe virus* (GRGV), which with GFkV constitute the „flex complex“ (Martelli, 2014).

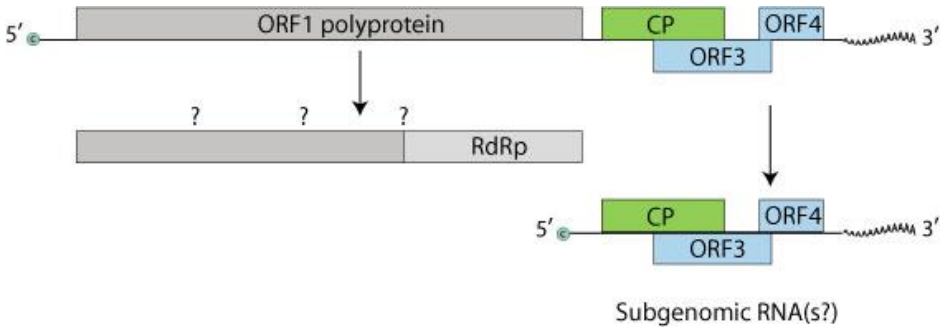



**Figure 5. Maculavirus genome**

**(Taken from: https://viralzone.expasy.org/56?outline=all_by_species)**

## 2.7 Fleck complex

The grapevine fleck complex is consisted of several diseases, and most of them can be detected by grafting onto the indicator *Vitis rupestris* St. George (Meng et al., 2017).

When it comes to nomenclature, current classification assigns GFkV to the homonymous type species of the genus Maculavirus, other GFkV-like viruses have not been officially classified since there is a lack of complete genome sequences, although, taxonomic position of these viruses based on the available information on particle morphology, physicochemical properties, partial genome sequences, parwise comparisons and phylogenetic analyses (Meng et al., 2017).

Even though there are differences in the organization of the different GFkV-like virus genomes, all of them have several features in common which are following:

1. They are made up of a single molecule of positive-sense single-stranded RNA
2. They are polyadenylated at the 3' end
3. Pressumably capped at the 5' terminus
4. They are rich in cytidine (its content is over 40%)
5. Expressed via a combination of posttranslational processing of a large precursor polyprotein into several mature proteins included in viral replication and synthesis of 3' coterminal subgenomic RNA molecules as templates for CP translation (Meng et al., 2017).

Grapevine rupestris vein feathering virus (GRVFV)

Genome size of GRVFV is 6.7 kb (Figure 6) and is monocistronic and closely resembles to oat blue dwarf virus regarding its organization (Meng et al., 2017). The large ORF codes for a putative polyprotein of approximately 234 kDa and it contains conserved motifs of proteins involved in replication of virus, and possibly, two CPs with estimated molecular masses of 23 kDa and 21 kDa.


**Figure 6. Genome of GRVFV**
**(Taken from: Meng et al., 2017)**

## 2.8 Methods for sampling and detecting plant viruses

### 2.8.1 ELISA

Enzyme-linked immunosorbent assay, commonly known as ELISA is an immunological assay used to measure antigens, antibodies, proteins and glycoproteins in samples. It can be used for detection of the abovementioned components in different biological samples, humans, animals and plants.

Several variations of this test can be devised. Selection depends on the sensitivity, specificity, presence of interfering factors and the type and activities of antisera available. Variations are following:

1. Double antibody sandwich (DAS)
2. Double antibody sandwich indirect (DAS-I)
3. Plate-trapped antigen
4. PAS-ELISA

### 2.8.2 Double antibody sandwich

This method is the one that has been the most commonly used for plant virus detection since its description by Clark and Adams in 1977.

The imunosorbent surface is a plastic microtitre plate with wells that are designed for ELISA. Second step is adding a dilute solution of unlabelled antibody to the wells of the plate and the antibody adsorbed on the plastic becomes to the trapping antibody, TA. Following step is washing to remove any excess antibody and after, the sample (antigen) is added. Antigens specific to the bound trapping antibody attach themselves to it, but the rest of the proteins remain in the solution and are afterwards removed by washing. The antigen attached to the trapping antibody is detected by adding a labelled antibody, which is specific to the antigen. The label is the enzyme that has been previously conjugated to the antibody. The final step is addition of the

specific colorimetric substrate for the conjugated enzyme which leads to the development of colour as a result of enzyme action. The amount of colour and the rate of its development are in correlation to the amount of labelled antibody bound to the antigen which had been trapped by the antibody attached to the plate.

### 2.8.3 Double antibody sandwich indirect (DAS-I)

Previously mentioned method can be converted into an indirect procedure; the first two steps are the same for both of these methods. However, in DAS-I, the antigen bound to the trapping antibody is detected by an unlabelled intermediate antibody (IA) which is specific to the same antigen but originates from an animal species different from the one used to prepare the trapping body. The unlabelled IA which attaches to the antigen is detected by an enzyme-labelled antibody (LA) specific to the IA. Since the IA is from a different species than the TA, the LA binds only to the IA and non-specific binding of the LA to the TA happens. Measurement of the amount of LA is done by adding substrate and measuring colour change as in DAS.

This method involves an additional step, but it is more sensitive and allows use of a commercially prepared enzyme-labelled antibody to the IA. Single LA can also be used for multiple virus detection systems.

### 2.8.4 Plate-trapped antigen

Approach of this method of ELISA is based around trapping the antigen on the plastic surface, then react the trapped antigen with an unlabelled intermediate antibody (IA) specific to it. The IA is then detected as in DAS-I with usage of an enzyme-labelled antibody (LA), which is specific to the IA. This technique is relatively simple and no advanced purification of antisera or conjugate preparation is involved, if a commercially prepared enzyme-labelled antibody to the unlabelled IA is used. It is usually less sensitive than DAS or DAS-I for use with crude plant extracts and may not be effective when concentration of antigen in the sample is low.

### 2.8.5 PAS-ELISA

Another form of indirect ELISA has been devised for detecting viruses in plants. This method uses protein A in two applications to sandwich antibody-antigen-antibody layers (Edwards and Cooper, 1985). The first applied layer of protein A prepares the plate for antibody layer coating. The second layer of protein A is conjugated to the enzyme and detects the second antibody layer. With usage of seven antisera, protein A sandwich ELISA (PAS-ELISA) was able to detect homologous virus isolates in standard dilutions of infected plant homogenates at values of A405 which were at least one absorbance unit grater than those of healthy controls. This method is more sensitive than the direct double antibody sandwich of DAS-ELISA.

### 2.8.6 Reverse transcription PCR (RT-PCR)

RT-PCR is technology that converts RNA molecules into their complementary DNA (cDNA), followed by amplification of the newly synthesized cDNA by standard PCR procedures. RT-PCR is a two-step process. Its mechanism involves reverse transcription of purified RNA by reverse transcriptase enzyme via an appropriate method for priming and amplification of first strand of cDNA. It is the most sensitive technique for detecting mRNA.

Since this technique is sensitive, specific and also inexpensive compared to many serological methods, and more reliable than them (Lopez et al., 2008), it has been developed and employed for detection of many plant viruses, for example, potato viruses such as PVX, PLRV, PVS, in stem or seeds of potato (Peter et al., 2009).

In research conducted by Gambino et al. (2006), multiplex reverse transcription-polymerase chain reaction (mRT-PCR) was developed in order to simultaneously detect nine grapevine viruses: *Arabis mosaic virus, Grapevine fanleaf virus, Grapevine virus A, Grapevine virus B, Rupestris stem pitting-associated virus, Grapevine fleck virus, Grapevine leafroll-associated viru*s -1, -2, -3. In the abovementioned research, mRT-PCR was able to give a reliable and rapid method for detection of grapevine viruses from a large number of samples.

### 2.8.7 Next generation sequencing

Next generation sequencing (NGS), also known as massively parallel or deep sequencing describing DNA sequencing technology which has revolutionised genomic research. Using this method, an entire human genome can be sequenced within a single day (Behjati and Tarpey, 2013).

There are a number of different NGS platforms using different sequencing technologies, but all of them have one thing in common; they all perform sequencing of millions of small fragments of DNA in parallel. Furthermore, bioinformatic analyses are used to combine together these fragments by mapping the individual reads to the human reference genome, or any other genome, depending on the type of the organism that is being researched. NGS can be used to sequence entire genomes or it can be used for investigating specific areas of interesting, whole exome or a small number of individual genes (Behjati and Tarpey, 2013).

When it comes to potential of NGS in clinical practice, there are several fields and advantages of using this technique:

1. It captures a broader spectrum of mutations than Sanger sequencing
2. Genomes can be interrogated without bias
3. Higher sensitivity of NGS allows detection of mosaic mutations

When it comes to disadvantages of NGS, they include putting in place the required infrastructure and the required expertise needed for analysing and interpreting collected data (Behjati and Tarpey, 2013).

Since the field of plant virology is in need of more sensitive techniques that would increase sensitivity, reliability and specificity of conventional PCR methods, several novel technologies have been developed and introduced during the recent years and in addition to having increased sensitivity, most of these methods also provide us with possibility of simultaneous detection of multiple pathogens (Czotter et al., 2015).

Some of the methods are: Loop-mediated isothermal amplification (LAMP), Micro and macroarray techniques, Barcodes and Deep (Next generation) sequencing.

Loop-mediated isothermal modification

LAMP was originally developed by Notomi and coworkers (2000) and it uses a set of inner and outer primers and Bst polymerase that amplifies DNA at 65$^{o}$C (Czotter et al., 2015). The elongation is followed by a colourimetric reaction with usage of, for example, hydroxynapthol blue (Goto et al., 2009). For this reason, using the thermal cycler, gel electrophoresis separation and ethidium bromide staining, even DNA exctraction, are not necessary to be able to score the results. For grapevine pathogens, this protocol has so far been used for the detection of *Xylella fastidiosa* (Harper et al., 2010) and also phytoplasmas (Tomlinson et al., 2010, Kogovšek et al., 2015).

### 2.8.8 Micro and macroarray techniques

The microarray and macroarray techniques involve the hybridization of PCR amplified and labelled samples to immobilized oligonucleotide probes that are specific for the virus pathogens (Czotter et al., 2015). Since microarrays allow the detection of ten-thousands of specific sequences in a single step of hybridization, this technique allows the detection of mixed infections or the complete virus population present in the tested plant (Czotter et al., 2015).

### 2.8.9 Barcodes

Barcoding is a technique that combines PCR amplification and sequencing. One of the basic conditions is that a barcode should contain highly conserved regions for given taxonomic group but sequences bordered by these regions should be variable enough to discriminate species or strains (Czotter et al., 2015).

### 2.8.10 Deep (Next generation) sequencing

Next generation sequencing is used as a base of finding all of the RNA that is expressed, including the pathogens RNA, in the host and get a virome of the plant. Deep sequencing gives a unique chance to get insight into any viruses or viroids present in the sample, whether they were expected or not (Czotter et al., 2015). So far, different platforms have been used for the description of new grapevine viruses; for example: Roche 454 for Grapevine Syrah-1 virus F (Al Rwahnih et al., 2012, Al Rwahnih et al., 2013).

## 3. MATERIALS AND METHODS

Samples of interest for this thesis were gathered at a grapeyard near Pecs Polgar pince-PP, since the GRVFV was present at that location. 11 individual grape samples were collected and the extracted RNA was used as template for the RT-PCR reaction. The concentration and the purity of the RNA extracts were measured by NanoDrop spectrophotometer.

### 3.1 First strand cDNA synthesis

For cDNA synthesis Revert Aid$^{TM}$ First Strand cDNA Synthesis Kit (Thermo Fisher) was used. We used virus-specific primer for the reverse transcription to increase the concentration of the virus-specific cDNA.

GRVFV specific primer: GRGVFV-R/6391: 5'-GCGCATTTCRTGGTGGTGCCGG-3'

The first step of the cDNA synthesis was the denaturating procedure. The following components were added into each of 0.5 ml PCR tubes:

- 10 µM virus-specific primer: 0.25 µl
- template RNA: 0.5 µg
- sterile milliQ water: to 3.12 µl

The samples were incubated at 65ºC for 5 minutes then immediately were chilled on ice. The reaction mixture for each denaturated RNA was prepared by adding the reagents below:

- 5X Reaction buffer: 1 µl
- 10 mM dNTPs: 0.5 µl
- Ribololck RNase inhibitor: 0.13 µl
- Revert Aid reverse transcriptase enzyme: 0.25 µl

The incubation settings were 42ºC for 60 minutes and for terminate the reaction the samples were heated 70ºC for 5 minutes.

### 3.2 PCR

The PCR proceeded using by Q5 High-Fidelity DNA Polymerase enzyme and GRVFV specific primer pair designed by Czotter et al. (2018) which amplify 1413 bp long PCR product.

1. Forward primer: GRGVFV-F/3501 (5'- CCTGCTGATCGCTGGAGACTCG-3')
2. Reverse primer: GRGVFV-R/4914 (5'-CGAAGATTCGCTGGTACTTCTT-3')

Master mix contained 14.5 µl PCR mix + 0.5 µl template that was made by cDNA synthesis. The reaction mixture consisted of the following components:

1. 5X Q5 Reaction Buffer: 3 µl
2. 10 mM dNTPs: 0.3 µl
3. 10 µM forward primer: 0.8 µl
4. 10 µM reverse primer: 0.8 µl
5. Q5 High-Fidelity Polymerase: 0.15 µl
6. sterile milliQ water: 9.45 µl
7. Template cDNA: 0.5 µl

Conditions of the PCR are shown in Table 1.

| PCR Step | Temperature (ºC) | Duration |
|---|---|---|
| Initial denaturation | 98 ºC | 30s |
| Denaturation | 98 ºC | 10s |
| Annealing | 55 ºC | 20s |
| Elongation | 72 ºC | 30s |
| Extension | 72 ºC | 10min |
| Number of cycles | 40 | |

**Table 1. Conditions of PCR**

### 3.3 Agarose gel electrophoresis

For analyzing the DNA fragments amplified by RT-PCR, used technique was agarose gel electrophoresis.

Preparation of 1.2% agarose gel consists of measuring 3.6 g of agarose and mixing it with 300 ml of 1xTBE in an Erlenmeyer flask. Following this step was microwaving the solution, until the agarose is dissolved completely, and it takes appeoximately 1-3 minutes for this event to occur. Solution needs to be cooled down, and in the next step of preparation, total amount of 1µl of ethidium bromide (EtBr) was added to a volume of 25-30 ml of agarose gel. Final step is pouring the agarose onto a gel glass, adding the wells comb, and waiting around 10-15 minutes for complete solidification of the gel.

To separate samples on the 1.2% agarose gel DNA loading dye (containing Bromophenol Blue and glycerol in TBE buffer) was added to the samples (1:5). Samples were applied into the wells and they were separated at 110 V until the Bromophenol Blue reached the bottom of the gel. Results were visualized on UV light.

### 3.4 Purification of PCR fragments from agarose gel

For purifying the 1413 bp long PCR product of interest from the gel, Thermo Scientific GeneJet Gel Exctraction Kit was used. The procedure of the purification is necessary for direct sequencing of the PCR product or for the cloning of the PCR fragment.

First step of the procedure consists of excising a DNA product with a sterile scalpel form the agarose gel, with attention given to trying to cut very close to the DNA band so the excess of the agarose gel is avoided. After cutting, the selected piece of gel was placed in a 1.5 ml microcentrifuge tube, which had to be weighed previously. The tube, containing the gel, was now weighed again and a volume of 1:1 of binding buffer was added to the slice as to dissolve the agarose, cause denaturation of proteins and to promote binding of the DNA to the column. Following step was incubation at 60°C for 10 min, so the gel slice can be completely dissolved.

After its dissolving, the solubilized gel solution was transferred into the GeneJet purification column. It was centrifuged for 1 min with discarding of the flow-through. In the next step, 100 µl of Binding buffer was added to the purification column, followed by centrifugation for 1 min, and discarding the flow-through. Then, 700 µl of Wash Buffer was added to the column and centrifuged for another minute, followed by the discarding of the flow-through. After this, the empty GeneJET column was centrifuged for 1 min, in order to remove the residual wash buffer.

Final step comprised of placing the column into a new Eppendorf tube, adding 25 µl of elution buffer and centrifuging it for 1 min and storing the eluted DNA at -20°C until they were used for sequencing.

Results of the sequencing showed us that the samples contained mixture of GRVFV strains, which is why each of the products was cloned into a pJET1.2 (Thermo Scientific) vector.

### 3.5 Cloning

The process of cloning has been executed with the usage of the Thermo Scientific CloneJET PCR Cloning Kit. The cloning vector, pJET1.2/blunt, is able to accept inserts whose size is from 6 bp to 10 kb. 1.5 ml Eppendorf tube was used for making a ligation mixture from 7.5 µl 2X Reaction Buffer, 0.75 µl of the pJET1.2/blunt Cloning Vector, 0.5 µl of T4 DNA ligase, 1 µl of nuclease-free water and 5 µl of the DNA fragment. This ligation mix was then incubated at room temperature for 5 min after which, the ligation mix was ready for transformation.

### 3.6 Transformation

Competent cells of *E. coli* (DH5 alpha strain), stored at -70, were put on ice in order to defrost, in duration of 10 to 15 minutes. Empty transformation tubes were placed on ice to which 200 µl of competent cells were added, with 5 µl of the ligation mixture. The mixture was stored on ice for 20 minutes.

After this, the tubes were transferred to a water bath of 42°C for a heat shock, for 30 seconds and then the tubes were moved back to ice.

500 µl of SOC medium without antibiotics was added to the tubes, which were then transferred to a shaking incubator at 37˚C. Incubation of the cultures lasted 40 minutes, which allowed the bacteria to recover and express the antibiotic resistance marker encoded in the plasmid. When the incubation was done, the transformed competent cells, in volume of 250 µl, were transferred and spread carefully so that not to damage the medium, which consisted of LB and ampicillin. The plates were left to incubate overnight at 37°C and the appearance of transformed colonies was expected in 12-16 hours.

### 3.7 Inoculation of liquid culture

250 µl of ampicillin was added to an Erlenmeyer flask filled with 250 ml of LB medium. Then, 3ml of LB medium + ampicillin was transferred to the inoculation flasks. From previously incubated Petri dishes, 10 colonies from each dish were picked and collected with a toothpick which we used to make a line in another Petri plate, previously labeled according to the sample, and after, the toothpick was placed to a small flask containing the 3ml of the medium. These cultures were incubated in a shaker at 37°C overnight. From each cloning experiments we grew 10 individual colonies.

### 3.8 Plasmid purification

The purification of Plasmid DNA was done by using the NucleoSpin Plasmid kit (Macherey-Nagel). In two steps all of the 3 ml transformed *E.coli* cells was taken and transferred into a 1.5 ml Eppendorf tube and centrifuged for 3 minutes at 8000 rpm in order to sediment the cells.

The supernatant was discarded and the pellet was kept. We re-suspended the remaining pelleted cells with adding 250 µl of Buffer A1. Complete suspension of bacterial cells was made by pipetting the content of the tube up and down.

Then, 250 µl of Lysis Buffer A2 was added and mixed thoroughly by inverting the tube gently, 6 times, and then left to incubate for 5 minutes at room temperature, until the lysate was clear.

300 µl of Precipitation Buffer A3 was added afterwards and the tube was gently inverted again, until the mixture was homogenous. The mixture was centrifuged for 5 minutes, at room temperature, at 11 000 g.

After centrifugation, the column was placed in a collection tube with adding of 750 µl of supernatant, and then centrifuged for 1 minute at 11 000 g. the flow-through was discarded, and the column was placed back into the collection tube.

Then, as a wash step 600 µl of Buffer A4 was added to the column and centrifuged at 11 000 g for 1 minute. The empty column was then placed back into the collection tube, centrifuged at 11 000 g for 2 minutes in order to dry the silica membrane and the column was set into a new 1.5 micro-centrifuge tube with addition of 30 µl of Buffer AE to elute the plasmid, incubated for 1 minute, then centrifuged for 1 minute at 11 000 g.

### 3.9 Plasmid digestion

Last step of the entire process included checking if the plasmid contained the cloned DNA. For that, digestion of the plasmids was done with two restriction enzymes, Xho1 and Xba1. Mixture for digestion contained 2 µl of 10X Tango Yellow Buffer, 0.2 µl of Xho1, 0.4 µl of Xba1 enzymes and finally 4.4 µl of MQ water. From this digestion mixture we add 7 µl to, 3µl of the purified plasmid and left to incubate for 1 hour at 37°C. Results of the digestion was checked by gel electrophoresis, after which the plasmids that did contain the inserted DNA fragment were sent to sequencing.

### 3.10 Analyzing of the sequences

#### 3.10.1 Chromas 2.1

After the sequences have been recieved, for the primary analysis of the results, Chromas 2.6.5 was used. This program is a free trace viewer for simple DNA sequencing projects that do not require assembling of multiple sequences. Among many other available options, it provides us with the possibility of copying the sequence to the clipboard in FASTA format or plain text, which can further be used for analyzing in other programs, depending of the need of the research. For the purposes of this research, it has been used to give us insight to wether we got clear sequences, or a mixture of them (Figure 7), and in our case, the sequences were indeed a mixture, which instructed us to further analyze them with other analyzing tools.

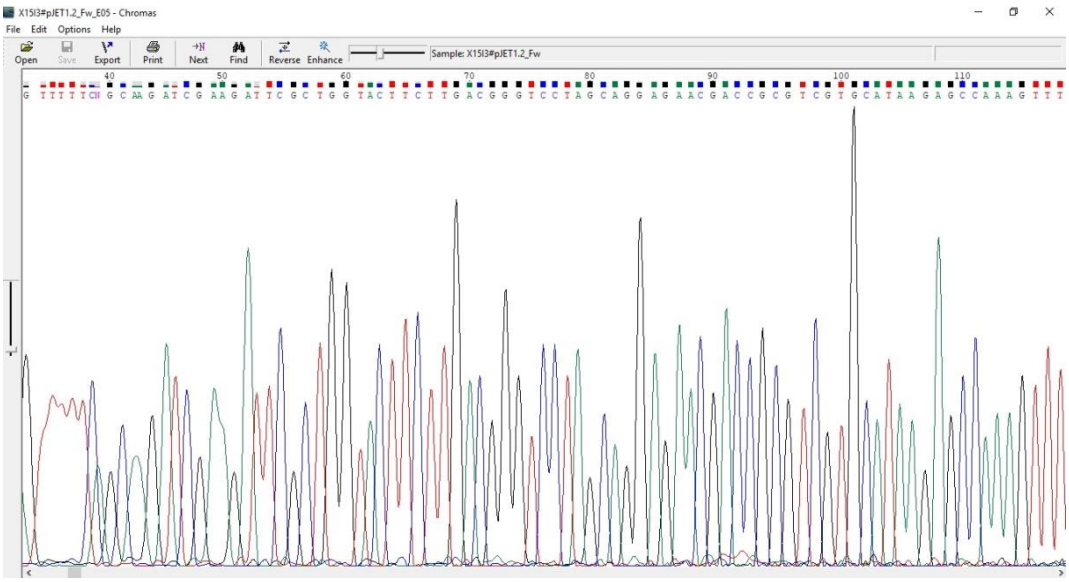



**Figure 7. Example of chromatogram of recieved sequence opened in Chromas 2.1**

### 3.10.2 Blastn

For further analyzation, Blastn from NCBI had to be used, in order to see the differences between the sequences and also to see the percentage of identity between the clones and to visualize the virus stretches. For this, Blastn has been used, with the option of multiple sequence alignment, so the sequences could be compared to one another and also to the reference genome of the GRVFV, AY706994. Also, it was optimized for „Somewhat similar sequences“ setting in order to get more valuable results for our type of research.

### 3.10.3 CLC Genomics Workbench

As a final tool, program called CLC Genomics Workbench has been used. It is a very powerful program for analyzing and visualizing data gathered from next generation sequencing (NGS). This program helped us to visualize comparisons done previously via Blastn and to compare results with each other so a final conclusion can be made. It was also used for creating a tree through Neighbor Joining method.

## 4. RESULTS

After the cDNA has been produced from every selected individual from the plantation and RT-PCR has been conducted, products that could be amplified were from samples: 2, 4, 5, 9 and 11 (Figure).

Products have been purified from the gel and sequenced by Sanger method with their forward primers. Results of this indicated that the samples were a mixture (Figure 8.) so each product was cloned into pJET1.2 vector.

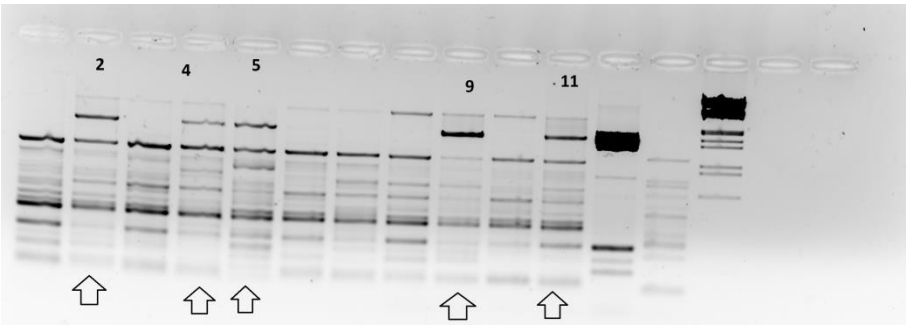



**Figure 8. Agarose gel electrophoresis that shows that products can be amplified from samples 2, 4, 5, 9 and 11**

Following this, purification of the products was made and the result of cloning was checked by cleavage with XhoI and Xbal restriction enzyme digestion. Results were:

1. 3 clones have been produced from sample X11/2 (X15/2, X15/3, X15/4) (Figure 9.)
2. 1 clone from X11/5 (X15/11)(Figure 10)
3. 4 clones from X11/9 (X15/13, X15/14, X15/15 and X15/16) (Figure 10)

Checking the results of cloning enabled us to collect necessary samples and send them to sequencing so the next step, that included bioinformatics work could be made and the sequences were aligned through several programs in order to come to a conclusion about GRVFV at our plantation.

After aligning the sequences in Blastn we have gathered information about the identity of the stretches of the cloned part of GRVFV. Figure 11 shows visual representation of Blastn alignment of the clones:

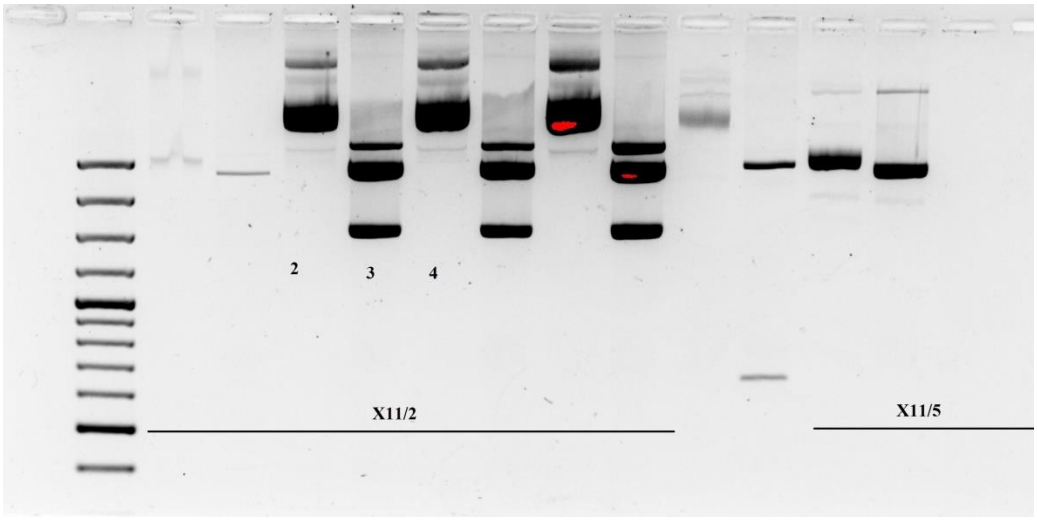



**Figure 9. Clones from X11/2**

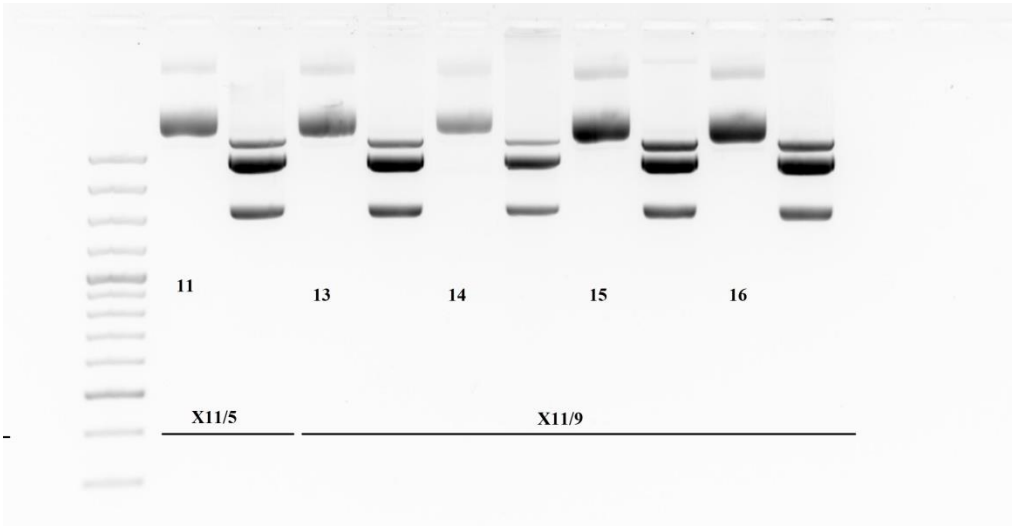



**Figure 10. Clones from X11/5 and X11/9**

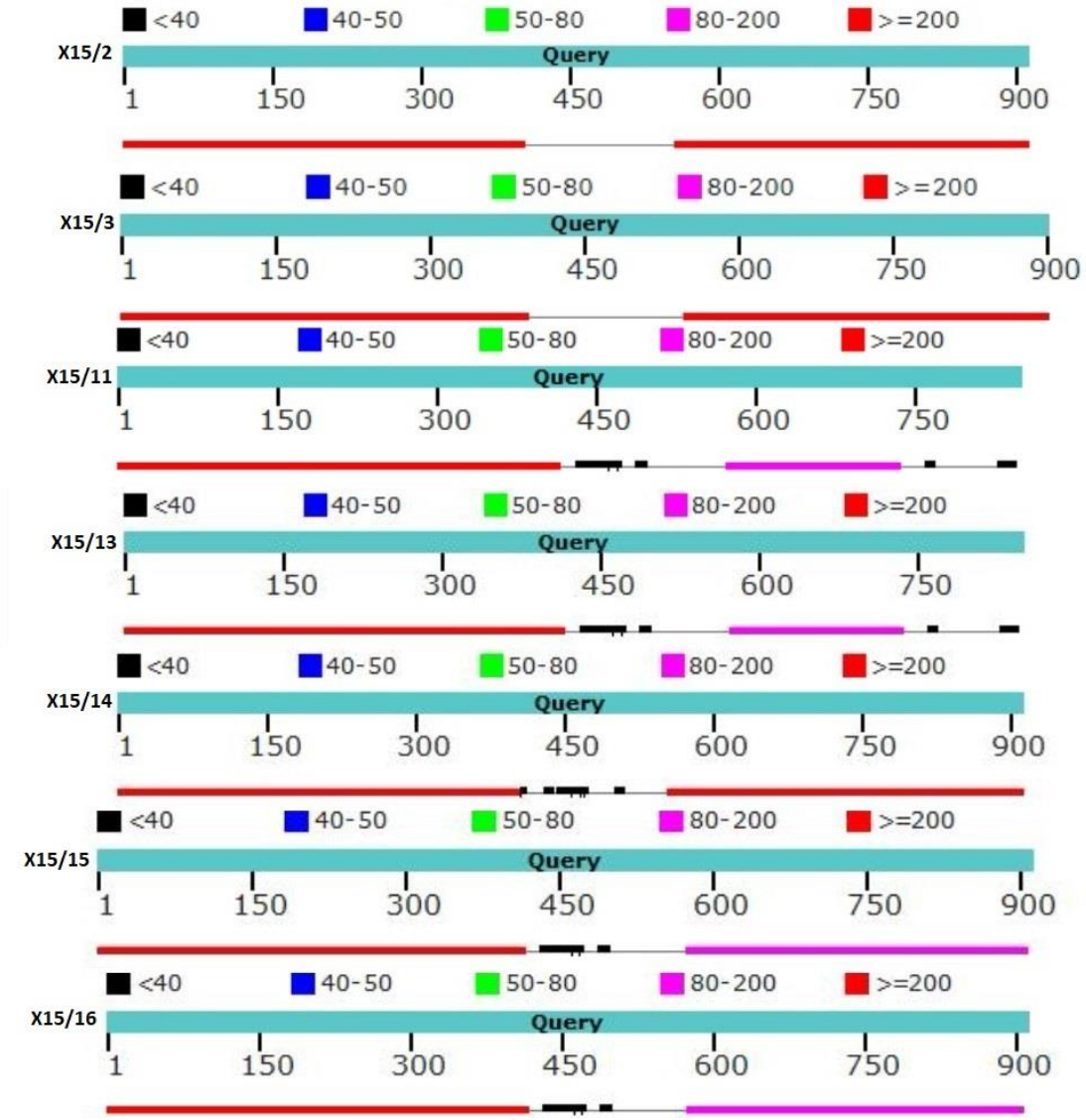



**Figure 11. Visual representation of Blastn sequence alignment**

Identity results are briefly summarized here as following:

Sequence X15/2:

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 432 bits(478) | 4e-124 | 338/404(84%) | 0/404(0%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 262 bits(290) | 5e-73 | 278/362(77%) | 12/362(3%) | Plus/Minus |

Sequence X15/3

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 407 bits(450) | 2e-11 | 6332/404(82%) | 7/404(1%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 268 bits(296) | 1e-74 | 277/360(77%) | 11/360(3%) | Plus/Minus |

Sequence X15/11

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 385 bits(200) | 6e-110 | 344/416(83%) | 0/416(0%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 173 bits(90) | 3e-46 | 140/165(85%) | 0/165(0%) | Plus/Minus |

Sequence X15/13

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 390 bits(203) | 1e-111 | 345/416(83%) | 0/416(0%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 167 bits(87) | 2e-44 | 139/165(84%) | 0/165(0%) | Plus/Minus |

Sequence X15/14

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 696 bits(362) | 0.0 | 390/404(97%) | 0/404(0%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 644 bits(335) | 0.0 | 351/359(98%) | 0/359(0%) | Plus/Minus |

Sequence X15/15

| **Score** | **Expect** | **Identities** | **Gaps** | **Strand** |
|---|---|---|---|---|
| 385 bits(200) | 6e-110 | 344/416(83%) | 0/416(0%) | Plus/Minus |

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 191 bits(99) | 2e-51 | 255/333(77%) | 0/333(0%) | Plus/Minus |

Sequence X15/16

| **Score** | **Expect** | **Identities** | **Gaps** | **Strand** |
|---|---|---|---|---|
| 385 bits(200) | 6e-110 | 344/416(83%) | 0/416(0%) | Plus/Minus |

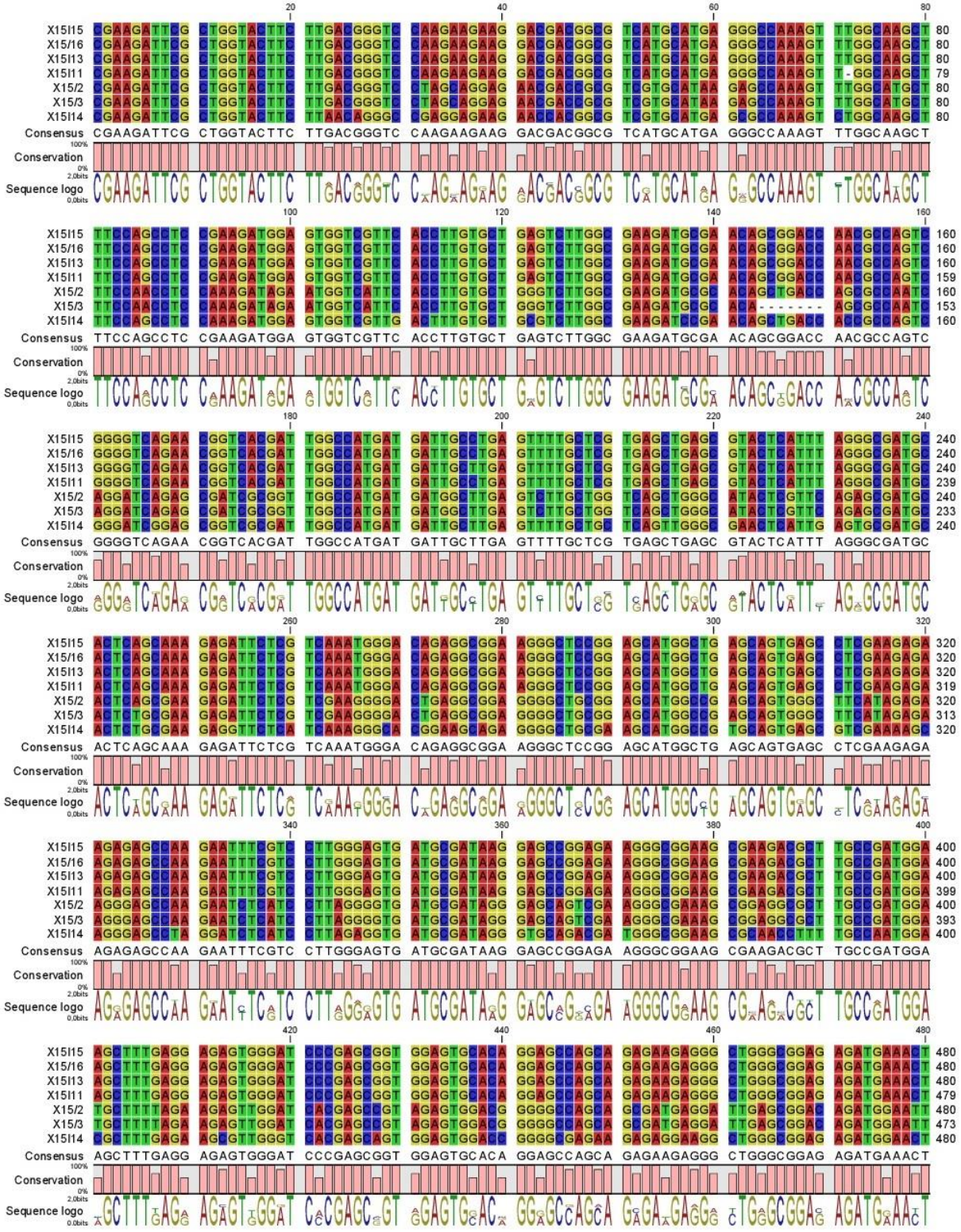
| **Score** | **Expect** | **Identities** | **Gaps** | **Strand** |
|---|---|---|---|---|
| 196 bits(102) | 3e-53 | 256/333(77%) | 0/333(0%) | Plus/Minus |

For easier interpretation of the results that were gathered through NCBI's Blastn, Table 2 summarize identities of stretches.

| Sequence name | Stretch 1 | Stretch 2 |
|---|---|---|
| | Identity (%) | |
| **X15/2** | 77 | 84 |
| **X15/3** | 77 | 82 |
| **X15/11** | 85 | 84 |
| **X15/13** | 84 | 83 |
| **X15/14** | 98 | 97 |
| **X15/15** | 77 | 83 |
| **X15/16** | 77 | 83 |

**Table 2. Identity of stretches between clones and GRVFV reference genome**

To compare all of the cloned GRVFV-s to each other a multiple alignment was produced using CLC Genomics Workbench in order to provide us with a deeper insight of the results (Figure 12). Moreover we have done phylogenetic analysis and produced phylogenetic trees using the same software package (Figure 13).


X15I15
X15/16
X15I13
X15I11
X15/2
X15/3
X15I14
Consensus
Conservation
Sequence logo

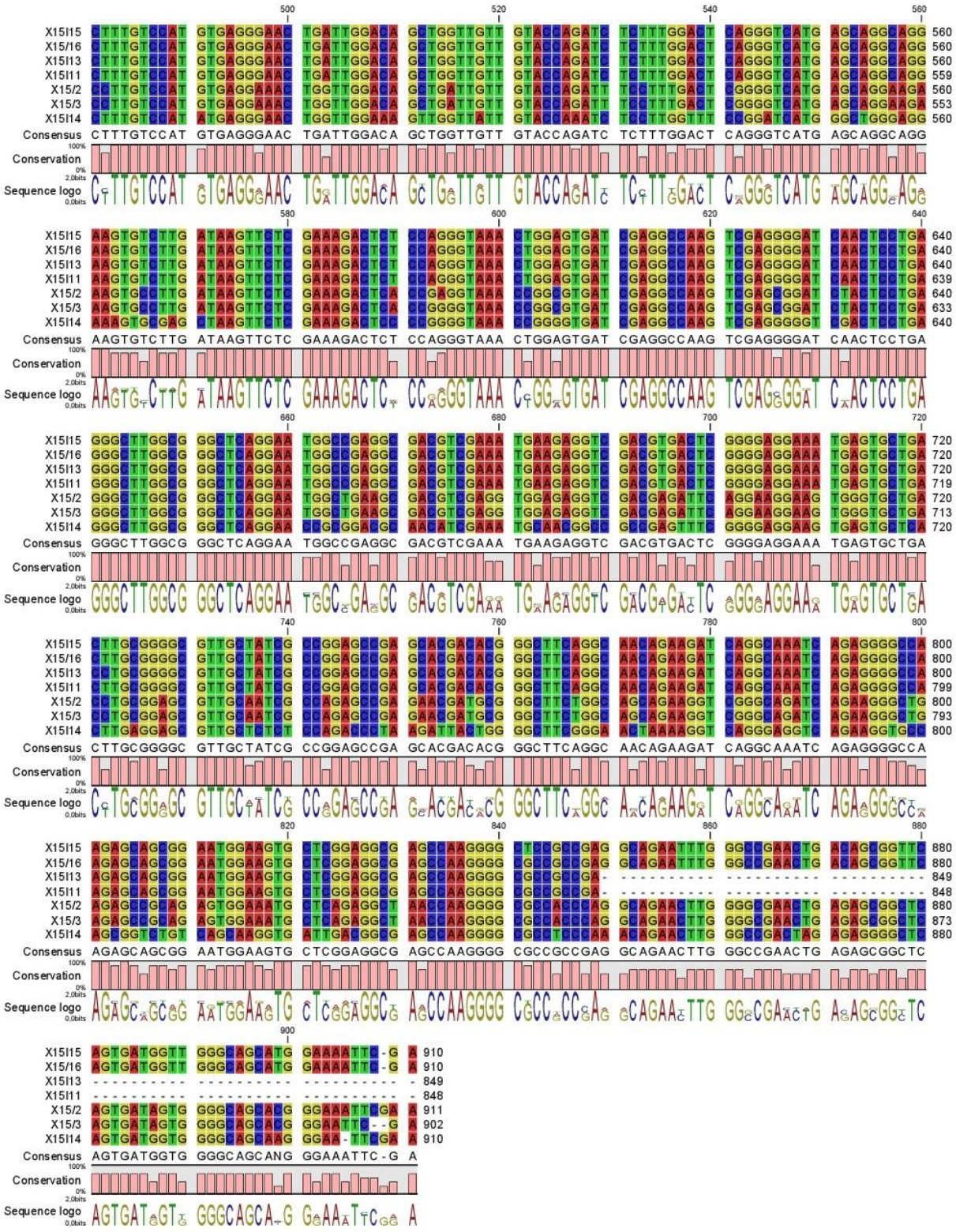



**Figure 12. Results of multiple alignment of cloned GRVFVs using CLC Genomics Workbench**

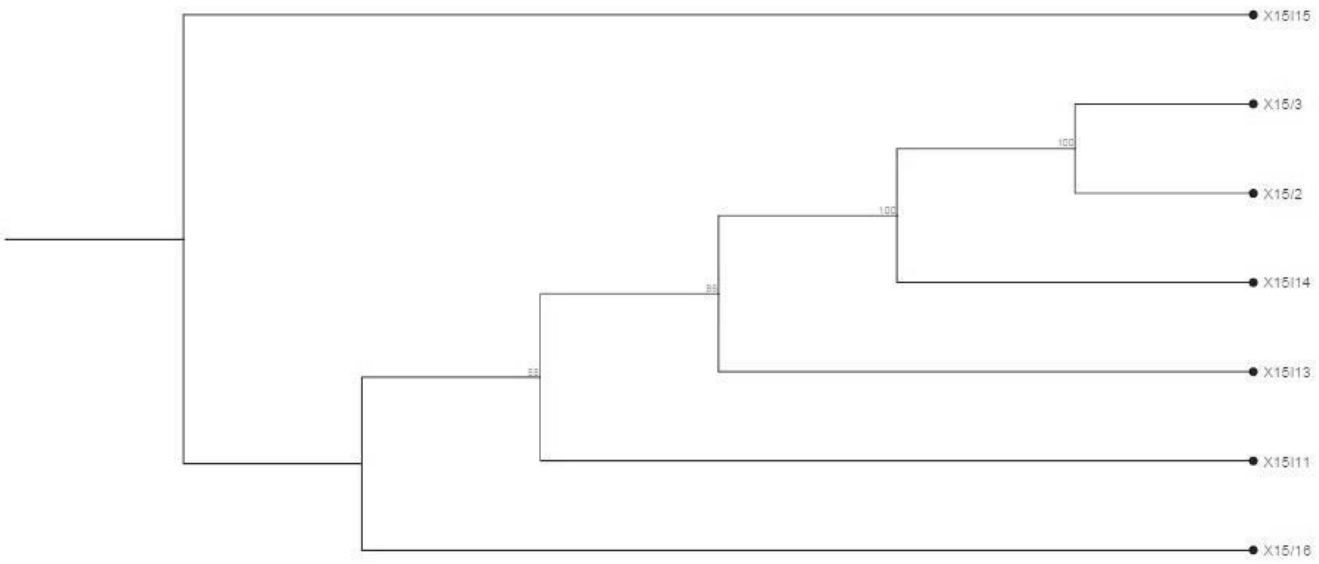



**Figure 13. Phylogenetic tree of cloned GRVFVs prepared by Neighbor Joining method using CLC Genomic workbench**

## 5. CONCLUSIONS

According to the above sequence analysis the following conclusions could be made:

1. Five grapevine individuals were infected by GRVFV in the investigated grapeyard
2. Cloning and sequencing more clones originated from one plant showed the presence of different virus variant not only at the grapeyard but also in all of the individuals.
3. In all GRVFV clones two regions show significant differences between one another when analyzed with Blastn.
4. The biggest differences between the two regions have been noticed in the X15/15 and X16/16 clones, originating from the same plant.
5. These findings were confirmed with the sequence analysis of the clones by CLC Genomics Workbench and show the existence of these differences which leads to a conclusion that a possible recombination happened not only at the plantation but in the same grapevine.
6. These separate recombinations at the individual level are the reason why there can be several different strains of the virus present within the pooled samples.
7. According to the phylogenetic tree it is obvious that clones originating from the same plant clustered distantly, however only 3 bootstrap values can be considered significant since they are over 70% identical, however in order to conclude more reliable conclusions further investigations are needed.

## 6. SUMMARY

**Thesis title:** Investigation of GRVFV variants in grapevine at a plantation near Pécs

**Author:** Dino Muratović

Course: Agricultural Biotechnology

Institute/Department: Institute of Genetics and Biotechnology

1. Primary thesis adviser: Dr. Erzsebet Kiss, Institute of Genetics and Biotechnology
2. Independent consultant: Dr. Éva Várallyay, *NAIK*, *MBK*, Diagnostic group

Grapevine (*Vitis spp*.) is one of the most widely grown of fruit crops worldwide. Recent worldwide production estimates (2014) are 7.6 million hectares and 74 milion metric tons. Grapevine finds its main uses for wine production, but also for fresh fruit, raisins, juice, vinegar, seed oils and several other products (Meng et. al, 2017).

Grapevine belongs to the group of the most cultivated horticultural crops which is a constant target of viruses due to its vegetative propagation. Since it belongs to woody plants, symptoms can go for a long period of time without being expressed and noticed, which is why it is important to regularly check the plantations in order to maintain virus free vineyard.

GRVFV was present at a grapeyard near Pecs Polgar pince–PP and from there 11 individual grape plants have been sampled. From this, later on, RNA was extracted and then pooled and this pool was used for amplification of GRVFV by RT-PCR.

From gathered GRVFV products, 2 clones have been sequenced from both sides, and these sequences, together with outher sequences of different variants have been deposited into GenBank.

In this research, cDNA has been produced from each individual at the plantation and RT-PCR has been done. Products that could be amplified were from samples: 2, 4, 5, 9 and 11.

These products were then purified from the gel and sequenced by Sanger method with their forward primers, but since they were shown as a mixture, each product was cloned into pJET vector. Following this was purification and the result of cloning was checked by cleavage with

XhoI and Xbal restriction enzyme digestion. According to this, 3 clones have been produced from sample X11/2 (X15/2, X15/3, X15/4), 1 clone from X11/5 (X15/11) and 4 clones from X11/9 (X15/13, X15/14, X15/15 and X15/16).

Further analysis of the sequences included alignments with Blastn on NCBI and with CLC Genomics Workbench, where also a tree was made using Neighbor Joining Method.

Gathered results showed that there are identity differences between selected sequences which lead us to the conclusion that there is recombination at the individual level, and not on the level of the entire population.

## 7. ACKNOWLEDGMENT

I am grateful to Dr. Várallyay Éva for her support and guidance throught my thesis work and for letting me do this research within *NAIK, MBK,* Diagnostic group. I am also very thankful to her PhD students who were always there to help when I needed extra guidance.

Your support and help was much appreciated!

## 8. LIST OF REFERENCES

### 8.1 Written sources:


1. Al Rwahnih M, Dave A, Anderson MM, Rowhani A, Uyemoto JK & Sudarshana MR (2013): Association of a DNA virus with grapevines affected by red blotch disease in California. *Phytopathology*. 103: 1069-1076
2. Al Rwahnih M, Shudarshana MR, Uyemoto JK & Rowhani A (2012): Complete genome sequence of a novel vitivirus isolated from grapevine. *Journal of Virology*. 86: 9545
3. Behjati, S., & Tarpey, P. S. (2013, August 28). What is next generation sequencing? Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3841808/
4. Czotter, N., Farkas, E., Lozsa, R., Ember, I., Szucsne, V., Varallyay, E., & Szeged, E. (2015). Primers designed for the detection of grapevine pathogens spreading with propagating material by quantitative real-time PCR. *International Journal of Horticultural Science*.
5. Edwards, M. L., & Cooper, J. I. (1985, August). Plant virus detection using a new form of indirect ELISA. Retrieved March 04, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/4055975
6. Fuchs M., Marsella-Herrick P., Loeb G. M., Martinson T. E., Hoch H. C. (2009). Diversity of ampeloviruses in mealybug and soft scale vectors and in grapevine hosts from leafroll-affected vineyards. *Phytopathology* 99: 1177–1184 [PubMed]
7. Galet, P. 2000. General Viticulture. (J. Towey, translator). Collection Avenir Oenologie. Oenoplurimedia, Chaintré, France.
8. Garnsey, S. M., & Cambra, M. (n.d.). Enzyme-linked immunosorbent assay (ELISA). Retrieved March 04, 2018, from http://www.fao.org/docrep/t0675e/T0675E0f.htm
9. Hewwitt, W.B., A.C Goheen, D.J. Raski, and G.V. Gooding Jr. 1962. Studies on virus diseases of the grapevine in California. Vitis 3: 57-83
10. Hewwitt, W.B., A.C Goheen, L. Cory, & C. Luhn. 1972. Grapevine fleck disease, latent in many varieties, is transmitted by graft inoculation. *Annales de Phytopathologie*, No. hors serie, 43-47
11. Hoefert L. L., Gifford E. M. , Jr (1967). Grapevine leafroll virus – history and anatomic effects.*Hilgardia* 38: 403–426

12. Lee, J., Cho, W., Lee, S., Choi, H., & Kim, K. (n.d.). Development of RT-PCR Based Method for Detecting Five Non-reported Quarantine Plant Viruses Infecting the Family Cucurbitaceae or Solanaceae. Retrieved March 06, 2018, from http://koreascience.or.kr/article/ArticleFullRecord.jsp?cn=E1PPBG_2011_v27n1_93
13. Lopez, M. M., Llop, P., Olmos, A., Marco-Noales, E., Cambra, M., & Bertolini, E. (2008, June 25). Are molecular tools solving the challenges posed by detection of plant pathogenic bacteria and viruses? Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18577779
14. Maliogka V. I., Dovas C. I., Lotos L., Efthimiou K., Katis N. I. (2009). Complete genome analysis and immunodetection of a member of a novel virus species belonging to the genus *Ampelovirus*.*Arch. Virol.* 154: 209–218 [PubMed]
15. Martelli, G.P. (2014). Directory of virus and virus-like diseases of the grapevine and their agents. *Journal of Plant Physiology* 96: 1-136.
16. Martelli, G.P., and C.E. Taylor. 1990. Distribution of viruses and their nematode vectors. *Advances in Disease Vector Research* 6: 151–189.
17. Martelli, G.P., J. Lehoczky, A. Quacquarelli, and G. Sarospataki. 1967. A disorder resembling “legno riccio” (rugose wood) of grapevine in Hungary. *Phytopathologia Mediterranea* 6: 110-112.
18. Meng, B., Martelli, G. P., Golino, D. A., & Fuchs, M. (2017). *Grapevine Viruses Molecular Biology, Diagnostics and Management*. Springer Verlag.
19. Over de Linden A. J., Chamberlain E. E. (1970). Effect of grapevine leafroll virus on vine growth and fruit yield and quality. *N. Z. J. Agric. Res.* 13: 689–698
20. Peter, K. A., Gildow, F., Palukaitis, P., & Gray, S. M. (2009, June). The C terminus of the polerovirus p5 readthrough domain limits virus infection to the phloem. Retrieved March 06, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/19297484
21. Rosciglione, B. & Gugerli, P. 1986. Maladies de l'enroulement et du bois strié de la vigne: analyse microscopique et sérologique. Rev. Suisse Vitic. Arboric. Hortic., 18: 207-211.
22. Savino, V., Boscia, D. & Martelli, G.P 1989. Rugose wood complex of grapevine: can grafting to Vitis indicators discriminate between diseases? Proc. 9th MEET. Icvg, Kiryat Anavim, Israel, 1987, p. 91-94
23. Scheu G. (1936). *Mein Winzerbuch*. Berlin: Reichsnährstand Verlags-Ges, 274 p

24. Thompson J. R., Fuchs M., Perry K. L. (2012). Genomic analysis of grapevine leafroll associated virus-5 and related viruses. *Virus Res.* 163 19–27 [PubMed]
25. Woodrum R. C., Antcliff A. J., Krake L. R., Taylor R. H. (1984). Yield differences between Sultana clones related to virus status and genetic factors. *Vitis* 23: 73–83
26. Zohary D, Hopf M (2000) Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe, and the Nile Valley, 3rd edn. Oxford University, New York.

**8.2 Internet-based sources:**


27. https://viralzone.expasy.org/300?outline=all_by_species
28. https://viralzone.expasy.org/285?outline=all_by_species
29. https://viralzone.expasy.org/270?outline=all_by_species
30. https://viralzone.expasy.org/56?outline=all_by_species

## **9**. **DECLARATION**

Signed below, ……., student of the Faculty of Agricultural and Environmental Sciences, Szent István University, at the BSc/MSc Course of ……. declare that I have used the cited and quoted literature in accordance with the relevant legal and ethical rules. I understand that the one-page-summary of my thesis will be uploaded on the website of the Faculty/Institute/Course.

Confidential data are presented in the thesis: yes no

Gödöllő, …..day ……month…….year

…….

Student's signature

As primary thesis adviser of the author of this thesis, I hereby declare that review of the thesis was done thoroughly; literature sources cited in the dissertation were used in accordance with the relevant legal and ethical rules. I hereby, approve the thesis for oral defense on Final Examination.

Confidential data are presented in the thesis: yes no

Gödöllő, …..day ……month…….year

…….

Signature of the Primary Thesis Adviser