UNIVERSITY OF SOUTH BOHEMIA IN CESKE BUDEJOVICE
FACULTY OF AGRICULTURE

Ph.D. Thesis

Molecular analysis and genetic identification of a new potyvirus and phytoplasma plasmids

Supervisors: Doc. RNDr. Karel Petrzik, CSc.
Prof. Ing. Vladislav Čurn, Ph.D.

Autor: Ing.Tatiana Sarkisova

Ceske Budejovice, 2013

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval/a samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to [v nezkrácené podobě – v úpravě vzniklé vypuštěním vyznačených částí archivovaných Zemědělskou fakultou JU] elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách.

Ceské Budejovice ....

Tatiana Sarkisova

Financial support

This work was supported by The Grant Agency of the Academy of Sciences of the Czech Republic [Grant AV0Z50510513] and by [Grant QH71145] of the National Agency for Agriculture Research of the Czech Republic.

Acknowledgements

I acknowledge my research advisors for the professional guidance and invaluable advice that helped me in reaching my research aims. I really appreciate their insight in resolving the numerous research challenges and for being very kind and understanding my needs.

I would like to thank my family, lab mates and close for me people. Without their patience, understanding, support and most of all love, the completion of this work would not have been possible.

Summary

Fabaceae or *Leguminosae* is known as a big family of flowering plants. It is the third biggest family after *Orchidaceae* and *Asteraceae*, which includes more than 19,400 species according (The Angiosperm Phylogeny Group, 1998). The current work was aimed primary for the screening symptomatic plant samples from family *Fabaceae* for the presence of viral and phytoplasma infection. As a result, new potyvirus - for which the name *Lupine mosaic virus*, LuMV was proposed - has been detected and analyzed.

The complete sequence of *Lupine mosaic virus,* LuMV was reconstructed using PCR with specific and degenerate potyvirus primers. PCR products were sequenced either directly or from cloned PCR products. Resulting sequence comprised of 10,113 nucleotides excluding the poly (A) tail and submitted to GenBank under accession number HM748648. The analysis of LuMV genome sequence showed four initiation codons within first 300 nucleotides of the long single open reading frame (ORF). However, only one was in a favorable context needed for translational initiation in plants (Lutcke et al.*,* 1987) and thus was the most likely the initiator of LuMV polyprotein in vivo. The UAA termination codon was located at 9899–9901 nt, followed by 212 nt of the 3'- untranslated region. Encoded large polyprotein was proteolytically cleaved into P1, helper component proteinase (HC-Pro), P3, cylindrical inclusion (CI) protein, 6K (6 kDa), viral genome-linked protein (VPg), nuclear inclusion proteins (NIa and NIb), and coat protein (CP) (Adams et al., 2005; Ng and Falk, 2006). The recently described putative protein PIPO (Chung et al., 2008) was identified within the region of P3 protein starting in the +2 ORF from a GAAA motive at position 3376. Predicted protein was 66 aa long with a (Mr.) of 7.3 kDa (Sarkisova et al., 2011).

The neighboring phylogenetic tree was created for the CP for distinguishing of the taxonomic status for viruses. *Lettuce mosaic virus,* LMV, *Plum pox virus*, PPV and *Panax virus Y*, PanVY were found to be the closest relatives. Amino acid sequence of *Lupine mosaic virus*, LuMV was found to be less than 80% thus creating a new potyvirus species, according to ICTV criteria for genus Potyvirus (Adams et al., 2005; Desselberger et al., 2009).

In 19 out of 37 samples, an extrachromosomal DNA from phytoplasmas has been detected. Phylogenetic analysis based on partial fragments of *rep* gene showed

several clades. Plasmids found in phytoplasma-infected particular 16S ribosomal group were not always in the same cluster, which can suggest that plasmids do not have close relationships with their phytoplasma genomes.

Two plasmids from plant-pathogenic mollicute "*Candidatus* Phytoplasma *Pichris echioides yellows,* PEY and *Psammotettix cephalotes' flower stunt phytoplasma*, BVK associated with phytoplasma classified to the 16Sr IX and 16SrXI-C subgroup, respectively, were subsequently taken to complete sequencing. However, this has not finished yet, thus only partial sequences are available today. They were submitted to GenBank under accessions numbers: KC545788 (pBVK *rep* gene), KC505535 (pPEY *rep* gene with conservative motives in N-terminal part), KC545789 (unknown gene and *ssb* proteins of pPEY), KC545790 (pBVK unknown protein, *ssb* and N-terminal part of another unknown protein).

Souhrn

Čeleď bobovitých (*Fabaceae*) představuje velkou skupinu kvetoucích rostlin. Množstvím zástupců je třetí největší čeledí - hned po vstavačovitých (*Orchideaceae*) a hvězdnicovitých (*Asteraceae*) - zahrnující 19 400 druhů (The Angiosperm Phylogeny Group, 1998). Předkládaná disertační práce se zabývá hledáním virů a fytoplazem způsobujících choroby zástupců této čeledi. Výsledkem je objev a analýza sekvence nového viru, pro který bylo navrženo jméno *Lupine mosaic virus*, LuMV. Zároveň byly zjištěny a molekulárně analyzovány fytoplazmové plazmidy, potvrzující infekci fytoplazmou.

Kompletní sekvence viru *Lupine mosaic virus*, LuMV byla poskládána z fragmentů získaných PCR pomocí specifických nebo degenerovaných primerů. PCR produkty byly sekvenovány buď přímo, nebo byly před sekvenováním klonovány. Výsledná sekvence sestávala z 10113 nukleotidů (bez poly-A konce) a byla uložena v GenBank pod přístupovým číslem HM748648. Její analýza identifikovala jeden dlouhý otevřený čtecí rámec (ORF) obsahující čtyři možné iniciační kodóny na úseku 300 nukleotidů. Pouze jeden z nich však leží v kontextu umožňujícím iniciaci translace u rostlin (Lutcke et al., 1987) a je tedy nejpravděpodobnějším startovním kodónem polyproteinu viru LuMV *in vivo*. Terminační kodón UAA leží na na pozici 9899-9901 a je následován 3'-netranslatovanou oblastí o délce 212 nt. Jediný dlouhý polyprotein je štěpen na protein P1, helper komponentu proteinázy (HC-Pro), protein P3, protein cylindrických inkluzí (CI), protein 6K, protein VPg (viral protein genome-linked), proteiny nukleárních inkluzí NIa a NIb a obalový (CP) protein (Adams et al., 2005; Ng and Falk, 2006). Uvnitř oblasti kódující protein P3 byl navíc nalezen nedávno popsaný předpokládaný protein PIPO (Chung et al., 2008). Jeho ORF začínala 2 nukleotidy před GAAA motivem na pozici 3376 a kódovala protein sestávající z 66ti aminokyselin o velikosti 7.3 kDa (Sarkisova et al., 2011).

Na základě aminokyselinových sekvencí obalového proteinu byl sestaven fylogenetický strom a nalezeni nejbližší příbuzní nově objeveného viru *Lupine mosaic virus,* LuMV: *Lettuce mosaic virus,* LMV, *Plum pox virus*,PPV a *Panax virus Y*, PanVY. Podobnost aminokyselinová sekvence polyproteinu LuMV a ostatních sekvenovaných potyvirů byla méně než z 80%. Podle současných pravidel ICTV pro rod *Potyvirus* (Adams et al., 2005; Desselberger et al., 2009) je tedy nově nalezený virus novým druhem.

V 19 ze 37 případů byla taktéž detekována přítomnost fytoplazmových plazmidů. Fylogenetická analýza částečných sekvencí *rep* genu odhalila několik skupin. Plazmidy stejné 16Sr skupiny fytoplazem ale nebyly vždy v jedné fylogenetické skupině. To může naznačovat, že plazmidy nejsou v úzkém spojení s jednotlivými genomy fytoplazem. Dva plazmidy fytoplazem Candidatus fytoplazma *Pichris echioides yellows*, PEY a *Psammotettix cephalotes' flower stunt fytoplazma*, BVK z 16Sr skupiny IX a XI-C byly vybrány pro kompletní sekvenování. Toto sekvenování dosud probíhá, proto jsou zatím k dispozici pouze částečné sekvence. Ty byly uloženy v GenBank pod přístupovými čísly: KC545788 (rep gen pBVK), KC505535 (rep gen pPEY s konzervovanými motivy na N-konci), KC545789 (neznámý protein a ssb pPEY) a KC545790 (neznámý protein pBVK, ssb a N-konec dalšího neznámého proteinu).

List of papers and author's contribution

Franova J., Petrzik K., Jakesova H., Beckova M., Sarkisova T. 2008. Diagnostic of Viruses infecting clover and Lucerne breeding material in the Czech Republic. Book of abstracts: the 3d conference of the international working group on Legume and vegetable viruses (IWGLVV):45.

T.Sarkisova contributed on electron microscopy observation of samples, nucleic acid extraction and manipulation, detection of viruses by amplification with group-specific primers, manuscript revision

Franova J., Petrzik K., Jakesova H., Beckova M., Sarkisova T. 2009. Cultivated and wild growing forage crops –reservoirs of viruses and phytoplasma. Grassland Sci. Eur. 14: 106-108

T.Sarkisova contributed on electron microscopy observation of samples, nucleic acid extraction and manipulation, detection of viruses by amplification with group-specific primers, manuscript revision

Sarkisova T., Petrzik K. 2009. Analysis of capsid protein sequence revealed new potyvirus found in *Lupinus polyphyllus* in the Czech Republic. Book of abstracts: XVIII Czech and Slovak plant protection conference. 35.

T.Sarkisova contributed on electron microscopy sample preparation and observation, performance of biological tests, nucleic acid extraction, primer design, capsid protein gene amplification and sequencing, manuscript writing and oral presentation on the conference.

Sarkisova T., Petrzik K. 2009. A new potyvirus identified in Czech Republic. Acta Virol.53: 143. **(IF: 0,682)**

T.Sarkisova contributed on electron microscopy observation, she was responsible for nucleic acid extraction and manipulation, amplification of the capsid protein gene, sequence analysis, manuscript writing.

Sarkisova T., Petrzik K. 2011. Determination of the complete nucleotide sequence of a lupine potyvirus isolate from Czech Republic reveals that it belongs to a new member of the genus Potyvirus. Arch. Virol. 156: 167-169. **(IF: 2,111)**

T.Sarkisova was responsible for biological tests of the virus, walking primer design, isolation of RNA, cDNA amplification, cloning, sequencing, contig assembling, correction of the complete sequence, manuscript writing and revision of the manuscript.

Petrzik K., Sarkisova T., Curnova L. 2011. Universal primers for plasmid detection and method for their relative quantification in phytoplasma-infected plants. Bull. Insect. 64: S25-S26. **(IF: 0,592)**

T.Sarkisova contributed on primer design, DNA isolation, and revision of the manuscript.

Lenz O., Markova J., Sarkisova T. 2011. Discriminating 16Sr groups of phytoplasma by an oligonucleotide microarray targeting 16S-23S ribosomal spacer. Bull. Insect. 64: 31-32. **(IF: 0,592)**

T.Sarkisova contributed on DNA isolation, and revision of the manuscript, she certified all samples used in the microarray design by specific amplification.

Sarkisova T., Petrzik K. 2012. Sequence analysis of two plasmids from *Pichris echioides yellows* and BVK phytoplasma from *Psammotettix cephalotes*. Book of abstracts: XIX Czech and Slovak plant protection conference. 154-155. ISBN 978-80-552-0838-1

T.Sarkisova was responsible for DNA isolation and management, primers proposal and selection, amplification, cloning, sequencing and sequence analysis, writing the manuscript and oral presentation on the conference.

Koloniuk I., Sarkisova T., Petrzik K. 2012. Evaluating detection methods of *Tolypocladium cylindrosporum* Mycovirus 1. XIX Czech and Slovak plant protection conference. 146. ISBN 978-80-552-0838-1

T.Sarkisova was responsible for fungi cultivation, microscopic identification of fungi and revision of the manuscript.

Co-authors confirmation:

…………………………………

Doc. RNDr. Karel Petrzik, CSc

…………………………………

Mgr. Ondrej Lenz, PhD

…………………………………

Ing. Jana Franova, PhD

Ceske Budejovice ____ 2013

…………………………………

Ing. Tatiana Sarkisova

Glossary

VLPs - Virus-Like-Particles

TEM - Transmission Electron Microscope

LuMV –Lupine mosaic virus

CP- capsid protein

P1 – first protein of the polyprotein precursor

HC-Pro - helper component proteinase

rep – replicase

P3- third protein of the potyviral polyprotein precursor

PIPO - pretty interesting potyviral ORF

6K- membrane associated protein

VPg- viral genome-linked protein

CI - cylindrical inclusion protein

3 - 5' UTRs - untranslated regions

ssb – single strand binding protein

pPEY- *Pichris echioides yellows* phytoplasma

pBVK - *Psammotettix cephalotes*' *flower stunt* phytoplasma

pOYW- plasmid from *Onion yellows* phytoplasma

Rap- replication-associated protein

pLS1- family members using rolling-circle replication strategy

ER- endoplasmic reticulum

aa- amino acids

nt- nucleotides

Content

General Introduction **14**

Objectives **16**

Chapter 1 .... **17**

1.1 Origins of viruses **17**

1.2 Virus evolution **18**

1.2.1 Mutation 18

1.2.2 Recombination and reassortment 19

1.3 Viruses on legume plants of family *Fabaceae* **20**

1.3.1 Family *Potyviridae* 25

1.3.1.1 Genus *Potyvirus* 25

1.4. Phytoplasmas and their plasmids **26**

1. 5 Material and methods **30**

1.5.1 Material 30

1.5.2 Enzymes and chemicals 31

1.5.2.2 Oligonucleotide (primers) 31

1.5.3 Equipment 33

1.5.4 Purification of Virus-Like Particles 33

1.5.5 Transmission Electron Microscope (TEM) 34

1.5.6. Phenol extraction method of total nucleic acids 34

1.5.7 Extraction of nucleic acids by using NucleoSpin® Extract II 34

1.5.8 cDNA production 34

1.5.9 cDNA with specific primers 35

1.5.10 PCR 35

1.5.11 Phi29 amplification or circular plasmid DNA 36

1.5.12 5´ RACE System for Rapid Amplification of cDNA Ends 36

1.5.13 Cloning and sequencing into pJET1.2 and TOPO vectors 36

1.5.14 DNA sequencing 37

1.5.15 Data analysis and accession numbers 37

Chapter 2 .... **38**

Results **38**

2.1. Molecular analysis of *Lupine mosaic virus*, LuMV – a new member in genus *Potyvirus* **38**

2.2. Analysis of LuMV RNA sequence **40**

2.2.1 Proteolytic processing of the polyprotein 42

2.2.2. Capsid protein (CP) 43

2.2.3 Replicase (*rep*) 45

2.2.3.1 Phylogenetic relationships 48

2.2.4 Role of helper component proteinase (HC-Pro) and capsid protein (CP) in potyviruses 51

2.2.5 P3 protein .... **Chyba! Záložka není definována.**

2.2.6 P1 protein .... **Chyba! Záložka není definována.**

2.2.7 Membrane associated protein (6K), viral genome-linked protein (VPg) and cylindrical inclusion (CI) protein 57

2.2.8 3 - 5' UTRs 61

Chapter 3 .... **68**

Results **68**

3.1 Summary of screening phytoplasma infected samples for the plasmid presence **68**

3.1.2 Detection of plasmids by using specific primers for *rep* gene 69

3.1.3 Partial sequence analysis of extrachromosomal DNA from *Pichris echioides yellows*, PEY and *Psammotettix cephalotes' flower stunt,* BVK phytoplasma 70

Chapter 4 .... **77**

General Discussion **77**

4.1 The complete nucleotide sequence of *Lupine mosaic virus*, LuMV - a new member in genus *Potyvirus* **77**

4.2 Screening samples for plasmids presence including partial sequenced **80**

pBVK and pPEY **80**

Conclusions **83**

References **84**

General Introduction

Fabaceae or *Leguminosae* is known as a big family of flowering plants. It is the third biggest family after *Orchidaceae* and *Asteraceae*, which includes more than 19,400 species according to Royal Botanical Gardens. Members of this family are found all over the world and grow in different environments and climatic zones (The Angiosperm Phylogeny Group, 1998). The most important species of agricultural plants are *Glycine max* (soya bean), *Phaselous* (bean), *Pisum sativum* (pea), *Medicago sativa* (alfalfa), and *Arachis hypogaea* (peanut), that are referred to well known members of this family. Species of weedy pests belonging to this family and growing in different parts of the world are *Cyticus* (broom) and *Pueraria lobata* (kudzu), and number of *Lupines* species (The angiosperm phylogeny group, 1998).

Viruses of plants make up a large and diverse group of pathogens. They can be a reason of cause serious diseases in different cultures and species of plants all round the world. Accurate identification of them is problematic because of a wide spectrum of plants-hosts and diversity of symptoms under different conditions. As early as 1960 significance and needs for standardized procedures for international identifying, including description and diagnosis of legume viruses, their classification, symptomatology, the environmental effect and method of transmission was pointed as necessity (Bos et al., 1960).

The characterization all viruses which have been found and described on plants from this family to present time and summarizing all results in just one review is impracticable desire. Thus, the characterization several important from distinct points have been talked in this review, some of them play the significant role as plant pathogens, which cause the large yield losses in agriculture important crops and infect herbaceous grasses.

The current work was aimed primary for the screening symptomatic plant samples from family *Fabaceae* for the presence of viral and phytoplasma infection. Plant specimens were with specific and degenerate primer's pairs to detect plant viruses belonging to following genera: *Potexirus, Potyvirus*, *Comovirus,* and *Carlavirus* and phytoplasma plasmids. For the detection, the reverse transcription and PCR were used with universal and degenerate specific primers.

During the routine detection, the sample of garden lupine was taken for detail analysis because of the suspicion it to be infected by virus and was subjected to

electron microscopy examination by which was confirmed a presence of filamentous virus particles of 690 nm in length. Sequencing of coat protein (CP) confirmed the presence of a new virus. Molecular characterization was aimed and the complete nucleotide sequence of the LuMV ssRNA was determined by amplifying and cloning of partial segments of the virus genome; reverse transcription and PCR, using degenerate and/or specific oligonucleotide primers has been done. A full genome sequence was submitted to GenBank database. Phylogenetic relatedness was evaluated by comparison with available data of other genus members.

Phytoplasmas are known quite well as a specific group of phytopathogenic organisms belonging to class *Mollicutes*; forming a pleomorphic group of bacteria because of lack a real cell wall and only surrounded by a three-layer membrane with a small and AT-rich genome. Their cell size is varied in diameters from 0.1 to 1 micrometer (Lee et al., 2000). Phytoplasmas are to be a reason to cause diseases in many plants and in some cases, yield loses, including economically important ones such as food, ornamental and fruit plants as was described previously (Lee et al., 2000; Seemüller et al., 1998).

In spite of the phytoplasma, which is, known since a long time ago the presence of extrachromosomal DNA was firstly reported about 25 years ago and there is little information about phytoplasma plasmids. Screening phytoplasma-infected samples for plasmids presence was aimed including evaluation of the phylogenetic relatedness between them. Search for conservative motives within *rep* gene was proposed to see its possible organization. Estimation of occurrence plasmids from different phytoplasma 16S ribosomal groups was planned in order to see their relationships with phytoplasma genomes.

Up to present, several phytoplasma plasmids have been already completely sequenced and submitted to GenBank database. However, there are a lot 'black spots' related to their functions and genome organization.

Objectives:

Consisting of following particular goals:

- Screen samples from family *Fabaceae* for the presence of viral and phytoplasma infection
- Screening will be done by molecular methods including specific and degenerative primers amplification and sequencing will be used for the detection of viruses and phytoplasma plasmids. These sequences data will be submitted to the GenBank
- Comparison and evaluation of obtained data. Multiply alignments and phylogenetic analysis will be performed to complete this particular goal
- *Rep* gene analysis for phytoplasma plasmids is proposed to know its organization. Search for conservative motives will be done along of the plasmid *rep* gene
- Estimation of plasmids occurrence from different phytoplasma 16S ribosomal groups
- Molecular characterization and sequencing of promising samples

Chapter 1

1.1 Origins of viruses

Unlike to other organisms, the origin of viruses and their evolution are still appeared to be unknown. However, there are three hypnotizes which explain their possible origin and evolution (Bubanovic et al., 2005). The cognition concerning their origin could assist the best understanding development of multi-cellular organisms' rapid diversification of species for period of time 600-700 million years (Bubanovic et al., 2005). Probably viruses have evolved based on natural selection pressure, similar to the other living beings and the high-degree variation of their genomes is provided by mutations and generic recombination (Domingo et al., 1997). Nowadays, in the age of molecular technology, new opportunities have been opened due to use of various molecular techniques and engineering, such as the polymerase chain reaction (PCR) and other methods that are applied in researches concerning viruses. It allowed us to enlarge our knowledge of viral genomes. Comparative analysis of the sequences of the virus genomes showed some similarities between viruses as well as between viral and cell proteins that probably can explain their conjectural origin and evolution (Gorbalenya, 1995; McGeoch, 1995).

Understanding the origin and evolution of viruses is not without difficulties, three current hypotheses are as follow: the first hypothesis is the theory of "regressive evolution," which implies that virus ancestors were free-living and more complex parasites. According to this theory, an increasing dependence of viruses on host-cell intracellular 'machinery' during evolutionary time was a crucial reason, meanwhile was retained the ability to auto-replicate, like mitochondria that have their own genetic information and replicate independently (Desjardins et al., 2005; Margulis et al., 2000). The second hypothesis, called the theory of 'cell origin', supposes that viruses originated of cell DNA and/or RNA, which acquired the ability to self-replicate, create extracellular virions, exist and function independently. The third one is the theory of 'independent' or 'parallel evolution of viruses and other organisms. According to it, viruses appeared at the same time as the most primitive organisms. The origin of the very first viruses can never be determined, however, they might have appeared long ago, over the

evolution of life on the Earth. It was supposed that the elementary form of life involving RNA replicons could be RNA viruses (Becker, 2000; Holland et al., 1998). RNA viruses might have continued evolving together with evolution of DNA world, because the cells, which had germinated DNA genomes, still contained the genes in their genome that coded RNA molecules and RNA polymerases. Sequence analysis confirmed this hypothesis and indicated that RNA and DNA viruses have common evolutional roots in their genomes and in some domains (Gorbalenya, 1995; McGeoch, 1995). The assumption was made that RNA viruses could arise in the nucleoprotein world owing to escape or reduction from RNA cells, whereas DNA viruses (at least some of them) could arise from RNA viruses directly (Forterre, 2006). Various mechanisms, such as mutations, recombination and re-assortment, which are the source of genetic variation, played greatly in their origin (Becker, 1998; Holland et al, 1998). Most probably or likely that viruses descended as result of many certain events, in which they have accumulated different genetic elements during the evolution of life (Holland et al, 1998).

1.2 Virus evolution

Three major strategies including recombination, mutation, and re-assortment could have created the necessary preconditions for evolution of viruses (Holland et al., 1998; Domingo et al., 1996). All these factors ensured and provided the diversity in viral genomes under natural selective pressure. Re-assortment is reputed to be considerable mechanism to change DNA. This model of genome reconstructs integral blocks of genes and transfers them to different and distinct locations which can be wherever in the genome as well as in genome of another host. One of the mechanisms of recombination is transduction, which is typical for both prokaryotic and eukaryotic organisms (Margulis et al., 1997; Margulis et al., 2000).

1.2.1 Mutation

RNA viruses use all possible and available mechanisms of genetic variation to provide their survival ability; all RNA viruses have extremely considerable rate of mutation due to lack of error proof activity of viral RNA-dependent RNA polymerase. A short period of replication time, a great amount of mutations and

mutation high-level are main properties of RNA virus replication. Because of all these factors RNA replicate as a dynamic and compound mass which is called viral quasi-species (Domingo et al., 1997). It is very important to estimate precisely mutation rate of viruses so that to understand them and struggle against them. Mutation rate of 23 viruses estimated as substitutions per nucleotide per cell infection (s/n/c). The results were within 10^{-8} -10^{-6} s/n/c (the mutation rate to substitutions per nucleotide per cell infection) for DNA viruses and within 10^{-6} - 10^{-4} s/n/c for RNA viruses, just as it had already showed for DNA viruses. In this case, negative correlation between mutation rate and size of genomes among RNA is obvious evident. However, additional experimental studies are required to confirm a statement like this. In contrast to some as opposite assumptions, mutation rate of retroviruses proved not to be lower as compared to other RNA viruses. It was also illustrated that number of point mutations is at average four times frequent as many as insertions/deletions (Sanjuan et al., 2010). High rate of mutations is able to generate RNA viruses with a great adaptive ability. In addition to the rate mutations, their frequency can be one of the reasons that form diversity in viruses (Domingo et al., 1997; Rossinck, 1997).

1.2.2 Recombination and reassortment

There are two types of genetic exchange to work in RNA viruses, they are recombination, and re-assortment as it has been already mentioned. Firstly, re-assortment is only found in multipartite viruses with substitution of one or more RNA molecules that make up the segmented viral genome. Secondly, the recombination process takes place in both types of viral genomes: segmented and monopartite. At this process donor of nucleotide sequence is introduced into the recipient RNA molecule and because of this is formation of new RNA containing genetic information from several sources (Worobey, 1999). Recombination is considered one of the main mechanisms in evolution of viruses. The recombination can occur due to homologous recombination between two nearly identical RNA or via non-homologous recombination between RNA (Simon et al., 1994). Phylogenetic analysis of many viruses - luteoviruses, nepoviruses and bromoviruses, showed that recombination could have played an important role in their evolution (Allison et al., 1989; Gibbs, 1995; LeGall et al., 1995).

1.3 Viruses on legume plants of family *Fabaceae*

Fabaceae or *Leguminosae* is known as a big family of flowering plants. Since 1970, approximately 100 articles are in the scientific journals on viruses infecting legumes in Canada such as alfalfa, bean, soybean, and sweet clover plants (Hamilton, 1997). A significant income to this research made the early reports on new viruses in Canada such as *Pea seed-borne mosaic,* PSbMV, *Lucerne transient streak,* LTSV, *Red clover necrotic mosaic*, RCNMV and *Sweet clover necrotic mosaic*, SCNMV. In addition, on distribution, cytopathology of viruses in seeds and pollen of *Alfalfa mosaic*, AMV and *Tobacco ringspot,* TobRSV and on detection of viruses on the pollen surface and possibility of transmission of viruses by such way (Hamilton, 1997). The very first mention of possible viral infection of legume in Canada was on bean crops in Prince Edward Island, where infections caused by *Alfalfa mosaic virus*, AMV were distinguished and unknown at that moment (Rankin et al., 1922). An interest to the evolution of viruses' dates from 1920, it was revealed that virus populations were heterogeneous and their structure was changing accordingly to experimental conditions. Before the advent of the molecular biology era, many significant data had been gathered about a host range and vector-associated transmission in the evolution of viruses as well as in small populations and co-adaptation of genes (García-Arenal et al., 2008).

There were reports about new viruses that had not been found in Canada before. One of them was *Pea seed-borne mosaic virus*, PSbMV transmitted by seeds and aphids. At the first time, potyvirus was revealed in the Northern America on the culture of peas in Wisconsin and Washington in 1968. Outbreak of the disease was observed in 1973 in California on USDA lines (Rao et al., 1985b). Later, *Clover primary leaf necrosis virus*, CPLNV was described, it was widely spread in Europe and Australia, and much later, it was designated as B serotype of RCNMV (Rao et al., 1987). *Lucerne transient streak virus*, LTSV, which was usually found in Australia, was registered in the Northern America at the first time, while researching viruses infecting legumes and already known by that time in Canada (Paliwal, 1982). Afterwards this virus was found on alfalfa in Alberta (Rao et al., 1985a). The SCNMV, S*weet clover necrotic mosaic virus* was found on yellow and white sweet clovers was classified and described, but previously this disease was recorded in Northern and Central Alberta in 1979, but at that time, it was unrecognized.

From year to year, the data on various viruses collected for last 20 years; many new results were obtained with a means of methods of molecular biology, the results were processed and described by Hamilton, (1997). The most important food crops are legumes such as faba bean, lentil, chickpea and pea and cereals such as bread and durum wheat and barley, that are cultivated everywhere especially in Central Asia and Northern Africa. Here they are the main source of carbohydrates and proteins for population (Makkouk et al., 2009). Infection contamination of these crops (cultures) occurs by natural way, by persistently transmitted aphid-borne, and number of viruses on these crops has been increasing, thereby inflicting a significant damage of legumes and cereals all over the world (Bos at el., 1988; Makkouk, 1994; Makkouk at el., 2003a; Kumari et al,, 2007). For the last three decades there have been carried out many surveys in the countries in West Asia and Northern Africa and the most significant and considerable viruses have been indicated. They are viruses that cause great damage to the plants grown there: *Faba bean necrotic yellows virus*, FBNYV, *Bean leafroll virus*, BLRV, *Beet western yellows virus*, BWYV, *Soybean dwarf virus*, SbDV and *Chickpea chlorotic stunt virus,* CpCSV , *Barley yellow dwarf virus*-PAV, BYDV-PAV, *Barley yellow dwarf virus*-MAV, BYDV-MAV and *Cereal yellow dwarf virus*-RPV, CYDV-RPV. If the infection starts at the beginning of the vegetation period, the yield losses caused by these viruses are very high (Loebenstein et al., 2004; Makkouk et al., 2009).

Accurate diagnosis together with sensitive rapid detection is extremely required for the effective management and control in legumes and cereals systems. The proper procedures of control could be introduced effectively only in a case of right diagnosed disease and if the area of its distribution is known. Over the last decades, there have been great achievements in increasing the sensitivities of the methods used for the detection of plant viruses (Makkouk et al., 2009). A major step forward was made with the advent and on coming of use of enzyme-linked immunosorbent assay (ELISA) which started use for the detection of plant viruses that replaced previously used serological methods, such as diffusion in the gel, especially in serial testing of samples (Clark et al., 1977). Afterwards it was improved and developed by monoclonal antibody technology and its application allowed identification of various viruses of legumes and cereals crops. Moreover, a different number of variants of ELISA was developed and introduced. That made possible to increase the diagnosis

accuracy and the sensitivity testing of many viruses of legumes and cereals crops (Makkouk et al., 1994; Makkouk et al., 1996). However, because of lack of necessary materials for the similar tests in many developing countries the tissue-blot immunoassay (TBIA) was developed and introduced, that enable to identify the most viruses of legumes and cereals crops (Makkouk et al., 1994; Makkouk et al., 1996). The RT-PCR was applied for detection of many viruses, which cause great damage, particularly five seed-borne legume viruses, and legume germplasm were identified. The following viruses are referred to them: *Alfalfa mosaic alfamovirus, AMV, Bean yellow mosaic potyvirus,* BYMV, *Clover yellow vein potyvirus,* ClYVV, *Cucumber mosaic cucumovirus,* CMV. All isolates of every virus were identified with the use of the RT-PCR assay that is five times sensitive as ELISA, which is alternatively more expensive and more time consuming (Bariana et al., 1994).

Identification and analysis of certain filamentous viruses were also carried out with the use of the degenerate primers in RT-PCR assay and *Closterovirus*, *Vitivirus* and *Trichovirus* genera were screening (Saldarelli et al., 2004). The analytical study of viral genomes conducted in 1980 and 1990 revealed quasi-species –like- structures in populations and enabled to analyze relationships among viral strains and species extensively. The huge size of the virus population and high level of adaptive mutations made up the concept, which becomes dominant in that period of time (García-Arenal et al., 2008). The most current classifications of viruses were developed on the base of capsid protein or polymerase gene sequences and their phylogenetic analysis. This is a more balanced approach that gives the total picture of relatedness by the whole genomes comparison (Stuart et al., 2004). Other features of viral genomes such as neutrality, multifunctionality of coded and non-coded sequences can limit the viral genome plasticity and affects their response to natural selection. Research of the viral evolution is still the issue of future, particularly due to its influence on host, insects, and dynamic of ecosystem (García-Arenal et al., 2008).

Molecular analysis is the base for classifications by genera, species, and strains. Certain various characteristics are usually used for discrimination of species among the genus. Criteria are not identical for all genera and families of viruses, data updated in ICTV reports (Van Regenmortel et al., 2000; Desselberger et al., 2009), and correspondent criteria are assigned for every genus. At the present differences

between sequences form very important element of these criteria in the most genera. Pairwise comparisons between sets of sequences have been used to define appropriate criteria for discriminating between species of the same virus, different species within the genus and different genus among potyviruses and geminiviruses (Fauquet, 2002; Fauquet, et al., 2003; Shukla, et al., 1994; Van Regenmortel, et al., 1997).

Identification and classification of potyviruses were unsatisfied because of a large group, a huge number of variations between members of the group and insufficient taxonomic parameters which would make it possible to distinguish all types of viruses from strains. Previously it used to be impossible to use terms “species” and “strain” to distinguish species of potyviruses from their strains, using methods such as host range and symptomatology, cross-protection, morphology of cytoplasmic inclusions and serology (Shukla and Ward, 1989). In contrast, based on nucleic acid and amino-acid sequences data it was shown that potyviruses could be divided in species and strains (Shukla et al., 1989).

Sequence data in combination with information about the structure of potyvirus particle were used to develop methods such as HPLC peptide profiling, cDNA hybridization. These findings along with data of immunochemical analyses created molecular basis for serology of potyviruses and explained many problems concerning with serological methods and became the basis for identification and classifications of potyviruses. Moreover, because of all these factors, virus/strain status of certain potyviruses was reconsidered and some changes were made in nomenclature. Under these conditions, all published data on symptomatology, cross-protection, and serology were required to be revised (Shukla et al., 1989). Analysis of the nucleotide and amino acid sequence data in the study of evolution and phylogeny was conducted; a 12 - nucleotide conserved sequence the ‘potybox’ was found which is considered to unique for the group of viruses that includes *Barley yellow mosaic virus*, BaYMV and transferred by pathogenic fungi. It was shown that various non-structural proteins of potyviruses have gemology with a completely unrelated viruses, at the same time 3'UTR and N'-terminal part of capsid protein are very variable among viruses, but are similar within species between strains. Hereby, they act as markers to estimate genetic relationships. By the present time, many

potyviruses have been sequenced and in all 3'- noncoding regions and the coat-protein are still taxonomic indicators in family *Potyviridae* (Atreya, 1992).

One of the sequenced potyviruses was *Bean common mosaic virus*, BCMV potyvirus. It is a pathogen of a common bean (*Phaselous vulgaris*) transferred by seeds (Barnett, 1991). Molecular analysis of whole genome was carried out and RNA consisted of 9612 nucleotides in length excluding 3´-terminal poly (A) tail with polyprotein 3066 aa long with a molecular mass (Mr.) of 310.3 kDa. Two viruses, *Bean common mosaic virus,* BCMV and *Bean common mosaic necrosis virus,* BCMNV are currently under control in Latin America and Africa Brown et al., (1990), reviewed the available data on geminiviruses of legumes in Latin America and the Caribbean. The isolates known as *Bean golden mosaic virus*, BGMV and other isolates such as *Bean calico mosaic virus*, BCMoV (Brown et al., 1990), and *Bean dwarf mosaic virus,* BDMV (Morales et al., 1990; Morales, et al., 1995) were studied too. It was concluded that all these viruses would have to be revised further again. The main host for *Bean golden mosaic virus,* BGMV is *Phaseolus vulgaris* and *Phaselous lunatus*. It was shown that different isolates found and they infected a number wild plants from family *Fabaceae*, especially, many viruses were found on such an undesirable plant as *Macroptilium lathyroides* and of course on species of *Phaseolus*, *Vigna* and *Calopogonium.* However, spectrum of plants-hosts is limited for BGMV in family *Fabaceae,* and *Malvastrum coromandelianum* was considered as a host. Golden mosaic symptoms had been found on legumes in many countries of tropical America and in many cases BGMV was specifically identified (EPPO/CABI, 1996). At the first time, this virus was discovered in 1976 in Columbia and caused huge losses of beans in Latin America. Losses amounted up to 75% in Brazil and this infection outbreak was directly associated with increasing of populations of *Bemisia tabaci* since 1970 (EPPO/CABI, 1996). This infection was still under constant attention because damage caused by it still not reducing. Particularly, if legumes are cultivated next to the sources of the vector, plants are infected at the early stage of plant development and disease spreads rapidly with the growth of plants. Moreover, increasing population of *B. tabaci,* biotype B promotes the early infection of leguminous plants (EPPO/CABI, 1996).

Viruses that were found on *Cassia bicapsularis, Voandezia subterranea,* and *Phaseolus lunatus* in Northern, West and East Tanzania and Kenya were serological

related with *Peanut mottle virus*, PMV. These viruses cause similar symptoms and spectrum of plants-hosts and differ by the virulence degree of certain species of plants (Saleh et al., 2005). Many viruses were identified on the plants of Vigna (cow pea) – this representative of tropical pulse plants, that plays considerable role in food production of developing countries in the tropics and subtropics, especially sub-Saharan Africa, Asia Central and South America (Kay, 1979). Vigna is affected by the plant pests' insects, and different pathogens, that attacked the plant at all stages of development (Allen, 1983). Crop losses caused by viral infection are within 10-100% dependently on interactions virus-host-vector, besides epidemiological factors also made significant contribution. More than 20 viruses were found on Vigna all over the world (Thottappilly et al., 1985; Mali et al., 1986; Brunt et al., 1990). About nine viruses – members of different genera were described in Nigeria (Thottappilly et al., 1992; Taiwo, 2003). *Cowpea mild mottle virus*, CPMMV was discovered on eggplants with mosaic symptoms *(Solanum melongena).* Filamentous particles with a normal length of 653 nm in *Leuconostoc lactis* were visible by electronic microscopy. Spectrum of plants-hosts of this virus is narrowly limited by members of plants from family *Solanaceae.* This virus is transferred by the aphid *Myzus persicae.* Higher infection with this virus was observed in autumn than in spring and *Solanum incanum* was recognized as a possible reservoir of CPMMV in summer time (Mansour et al., 2003).

1.3.1 Family *Potyviridae*

Viruses of the family *Potyviridae* infect plants and the family is composed of the six genera. The current taxonomic classification is entirely based on the VIIIth report of the International Committee on Taxonomy of Viruses (ICTV) (Desselberger et al., 2009). Family *Potyviridae* which is currently classified into six genera: *Ipomovirus, Macluravirus, Potyvirus, Rymovirus, Tritimovirus* – with monopartite and *Bymovirus,*with bipartite genomes is positive sense ssRNA; most of viruses are monopartite with the genome size 8 to 11 kbp, exception is genus *Bymovirus* with bipartite genome with 7.5 and 3.5 fragments (Shukla et al., 1998).

1.3.1.1 Genus *Potyvirus*

The potyviruses have filamentous particles about 700 nm in diameter built from CP protein. The infectious genome is presented by molecule positive sense ssRNA,

the 5′ end of each having a VPg and the 3′ end being polyadenylated. Functional viral proteins are produced by cleavages of the polyproteins by viral–encoded proteases at conserved processing sites. The host range of potyviruses is large. In addition, they can be readily transmitted by sap inoculation as well as by insect vectors: planthoppers and leafhoppers (Shukla et al., 1998).

1.4. Phytoplasmas and their plasmids

Phytoplasmas represent a specific group of phytopathogenic organisms belonging to class *Mollicutes*; forming a pleomorphic group of bacteria because of lack a real cell wall and only surrounded by a three-layer membrane with a small and AT-rich genome (Lee et al., 2000). Among the different plant distortions and abnormalities caused by phytoplasmas are dwarfing, phyllody of leaves and flowers, yellowing as well as flowers virescense, various tissue malformations on fruits and other growth disorders and aberrations. These symptoms were thought to cause by virus infections over a long period until Japanese scientists discovered morphological structures of phytoplasmas by electron microscopy. They are pleomorphic bodies because of absence cell wall remind structure of typical mycoplasmas by morphology and ultrastructure which are well known pathogens of humans and animals. These bacteria since many years had been called as mycoplasma-like organisms or MLOs because of found similarity in structure and organization. The complications which are related with their detection because of their inhabitation of phloem cells, concentration and no possible cultivation because all attempts to grow MLOs *in vitro* were unsuccessful (Doi et al., 1967; Seemüller et al., 1998).

Thus, for a long time the research related to taxonomy was stunned because of luck data for the definitive systematic and phytoplasmas were described and differentiated according to the symptoms they induce, the host plant affected, and sometimes also the geographic area where they occur, *e.g. European aster yellows* (Seemüller et al, 1998). The pathogen identification relied for more than 20 years on microscopic observations (DAPI staining) or electron microscopy detection. However, during last years the applications of DNA-based technology allowed to preliminary distinguish different molecular clusters of these prokaryotes (Bertaccini, 2007). Thus, a new step had been done only after coming era of molecular biology (Seemüller et al, 1998). DNA-based methods were introduced and allowed to begin a

new phytoplasma research, following the development of procedures to extract and enrich phytoplasmal DNA from infected plants or insects (Kirkpatrick et al., 1987; Lee et al., 1988; Sears et al., 1989; Kollar et al., 1990). The possibility to design specific primers for highly conserved genes such as 16S ribosomal gene together with the use of molecular probes randomly cloned from phytoplasma genome, allowed discriminating and molecularly classifying them. Now a certain amount of knowledge is available that allow starting epidemiological studies in order to prevent further spreading of phytoplasma-associated diseases (Bertaccini, 2007). Phytoplasmas DNA could be then amplified, cloned and sequenced and classified using the "Candidatus" concept where each of the major clades established by 16S rRNA sequence analysis represent a *Candidatus* species of the *Phytoplasma* genus (The IRPCM Phytoplasma/Spiroplasma Working Team—Phytoplasma taxonomy group, 2004).

A wealth of molecular data on phytoplasma diversity and on the relationships of the phytoplasma was generated. In several comprehensive studies on phylogeny and taxonomy of the phytoplasmas, many phytoplasmas from several phylogenetic groups have been examined, using either sequence or RFLP analysis of ribosomal DNA. Phytoplasmas for which 16S rDNA sequences are available have been classified into 20 major phylogenetic groups or subclades. Seventy-five phytoplasmas were distinguishable among the molecularly characterized phytopathogenic *Mollicutes* (Liefting et al., 2006; Seemüller et al, 1998) and up to present even more. Thus, molecular tools such as PCR/RFLP and nested-PCR of (16SrDNA) ribosomal phytoplasma region are developed and applied in order to get reliable system for phytoplasma detection and classification towards epidemiological studies of diseases associated with phytoplasma presence (Lee et al., 1998; Bertaccini, 2007). Development of polyclonal antisera first, and of monoclonal antisera later, allows to start first differentiations among phytoplasma groups (Bertaccini, 2007) while polyclonal antisera have relatively low specific titers, and are not readily useful for discrimination among phytoplasmas, the monoclonal antisera greatly improved the reliability of immune-identification techniques, such as ELISA, dot-blot immunoassays and immunofluorescence tests (Bertaccini, 2007).

The presence of extra chromosomal DNA, similar to plasmid DNA was demonstrated in phytoplasmas by using DNA probes; these DNAs (double stranded

covalently closed circle) could be different in different phytoplasma strains, but their role is still unknown in the majority of the cases (Bertaccini, 2007). Extrachromosomal DNA, including single-stranded (ss) and double-stranded DNA, associated with spiroplasma, mycoplasma, and Acholeplasma viruses has been described (Maniloff,1988; Razin,1985; Kuboyama et al.,1998). The *Mollicute* plasmids characterized to date are cryptic, and transfer of plasmid DNA between *Mollicute* has not yet been demonstrated (Kuske et al., 1990; Kuboyama et al., 1998; Nishigawa et al., 2001). However, plasmid DNA from bacteria has been shown to encode some biologically important genes, which could affect chemical tolerance, pathogenicity, virulence, and gene transfer (Davies et al., 1972; Panopoulos et al., 1985). Plant-pathogenic bacteria plasmids, like those in *Agrobacterium spp*., are essential for host-parasite interaction (Winans, 1992). In phytoplasmas, the biological functions of the plasmids and extrachromosomal DNAs have not yet been reported. The possibility that phytoplasmas may encode genes in the extrachromosomal DNA that are related to pathogenicity was shown (Kuboyama et al., 1998). Up to present several phytoplasma, plasmids have been completely sequenced in order to get more information about their structure - the genes their carry and proteins with unknown functions. As an instance, a plasmid was found in *Onion yellows* phytoplasmas strains and a 3.6-kbp DNA fragment was cloned from the extrachromosomal DNA of a pathogenic plant *Mollicute*, *Onion yellows* phytoplasma (OY-W). It was revealed by sequence analysis of the fragment that open reading frame (ORF) encoding the replication (*rep*) protein of rolling-circle replication (RCR)-type plasmids. This assumption was confirmed by detecting the single-stranded DNA (ssDNA) of a replication intermediate that is specifically produced by the RCR mechanism. This was the first report on the identification of the replication system of this plasmid and the genes encoded in it. In addition, by this work was also shown no homologues sequence between phytoplasmas genome and inhabited its plasmids and by this was concluded that plasmid has not been integrated to phytoplasma genome, as temperate phage would be (Kuboyama et al., 1998).

Comparative analysis of two plasmids of phytoplasma australiense has been done, ‘Candidatus Phytoplasma australiense’ is known to cause yellows/decline diseases of range plant hosts in Australia and New Zealand. The plasmids varied in their copy number and nucleotide sequence yet contained the same four open reading frames

(ORFs). The deduced amino acid sequence derived from ORF1 shared similarity with hypothetical proteins encoded on the plasmids from *Onion yellows* and *Beet leafhopper-transmitted virescence* agent phytoplasmas. The deduced amino acid sequences of both ORF2 and ORF3 share similarity with functionally unknown proteins on the chromosome of onion yellows phytoplasma (Liefting et al., 2006). During last decades, molecular characterization of several plasmids has been done. For instance, from paulownia witches'-broom phytoplasma was described and published a detection of plasmids. The two plasmids contained a series of tandem repeats and encoded a replication associated protein (*repA*) and a single-stranded DNA binding protein (*ssb*), which were necessary for the replication of plasmids. Seven putative proteins encoded by two plasmids were predicted to contain one or more hydrophobic transmembrane domains, respectively, and presumably to be localized to the membrane (Lin et al., 2009). Complete sequences for two plasmids associated with two strains of 'Candidatus Phytoplasma asteris' have been obtained. The plasmid named pPARG1 was found in *Rehmania glutinosa L*. associated with phytoplasma classified in the 16Sr I-C subgroup. Plasmid pPABN1 was from phytoplasma associated with infected winter oilseed rape and classified in the 16Sr I-B subgroup. The plasmids pPARG1 (4371 nt) and pPABN1 (3529 nt) have high A + T content of about 75%, similar to that of phytoplasma genomes (Petrzik et al., 2011).

Sequence analysis of two plasmids from the phytoplasma beet leafhopper-transmitted virescence agent was done and the complete nucleotide sequences of the two plasmids from, BLTVA have been determined. The larger plasmid, pBLTVA-1, was 10 785 nt in length and contained 11 putative ORFs, almost all of them were duplicated or triplicated on the plasmid due to the presence of large repeated regions. The sequence contained a series of tandem repeats, the largest of which was 338 nt long. The sequences of ORFs 4 and 11 showed homology with the replication genes of plasmids from other phytoplasmas and from geminiviruses. ORF9, the only ORF present as a single copy, showed homology with DNA primase genes from bacterial chromosomes and contained the conserved zinc finger and topoisomerase/primase domains. None of the other eight ORFs showed homology with known sequences in the GenBank database (Liefting et al., 2004).

The interesting results have been discussed in paper by Oshima et al., (2001). Research was performed by using plasmids from *Onion yellows* phytoplasma. Comparative sequencing analysis revealed this plasmid contains both plasmid and virus-like domains. C-terminal region was unexpectedly similar to the helicase domain of the replication-associated proteins (Rap) of eukaryotic viruses, especially circoviruses (ssDNA viruses of vertebrates). The extrachromosomal pOYW-*rep*, *Onion yellows* phytoplasma (OY-W), was specifically detected in OY-W-infected plant phloem cells by western blot, suggesting that it is a functional protein. The explanation of this suggested that an ancestral phytoplasma plasmid pOYW may have acquired a helicase domain from host phytoplasmal DNA, which entered the surrounding eukaryotic cytoplasm, and subsequently evolved into an ancestral eukaryotic ssDNA virus. Alternatively, a pOYW ancestor could have obtained the helicase domain by recombination with a virus: this would be then a first example of recombination between plasmids and viruses (Oshima et al., 2001).

Comparative analysis of all sequenced plasmids from various phytoplasmas indicated the diversity of phytoplasmal plasmids and despite of existence plasmids variations from different phytoplasmas the phylogenetic relationship of plasmids was consistent with the classification based on the 16S rDNA sequence of phytoplasmas (Liefting et al., 2004). The remarkable variation of plasmids in number and size was found in different phytoplasma groups or strains might be related to pathogenicity, vector transmission, or host adaptation. According to the roles of plasmids, which are known from other bacteria (Schneider et al., 1992; Chopra et al., 2001; Oshima et al., 2001; Vivian et al., 2001; Novichkov et al. 2004; Christensen et al. 2005; Wegrzyn, 2005; Lin et al., 2009).

1. 5 Material and methods

1.5.1 Material

Fresh samples of showing symptoms suspected for the virus presence were collected for the screening and among them, the garden lupine (*Lupinus polyphyllus*) showing mild mosaic symptoms and interveinal yellowing obtained from a private garden in the south of the Czech Republic. The suspected samples were maintained

in a glasshouse by mechanical inoculation to *Nicotiana benthamiana*, *Chenopodium quinoa* plants.

Frozen 37 samples from phytoplasma collection in Italy, Bologna were used in this work. They were given by Dr .Bertaccini A.

1.5.2 Enzymes and chemicals

T4 ligase, DNase, RNAase A, RNAase H, RevertAid™ MMLV Reverse Transcriptase, iScript^{TM} cDNA synthesis kit (Bio-Rad) and Monster Script^{TM} 1st-Strand cDNA Synthesis Kit (Epicentre, Technologies, USA), the DreamTaq polymerase (Fermentas, Lithuania). Chemicals used in the total RNA and DNA purification and buffers were obtained from Sigma-Aldrich (Germany), NucleoSpin RNA Plant Kit (Macherey-Nagel, Germany).

1.5.2.2 Oligonucleotide (primers)

The primers used in this study were designed using the DNASTAR software package (v.8.0.2)(Lasergene, Madison, WI, USA) and were synthesized by Sigma (Germany). The sequences of primers, listed in 5´-3´-direction, are shown in the table below:

Tab.1. List of primers used in this study:

5 ′AAP	**5'-GGCCACGCGTCGACTAGTACGGGIIGGGII GGG IIG -3'**
Spoty	**5'-GGNAAYAAYAGYCAR-3'**
LupCPr	**5'-GCATGTACGATCTTTCAGTATTTCTC-3'**
Lupr	**5'-GCTAACAGCAAATCGTCTCC-3'**
Lup N3	**5'-GCAYGTNGTNAARGGNAGATG-3'**
LupRE	**5 '-CATATACCAAGTTTGCCGAGG-3'**
Lup akvay	**5'-CGCTAAAGTTGCGTACACGC-3'**
up7000r	**5'-CGACGTCATCATGGTTGA-3'**
Lup	**5’-GTGGACCCATTAACTGGAGC-3'**
Lup6181r	**5'-ACGGAATGGATTGTGTGGTG-3'**
LuOAH	**5'-GCTTGCTTTGTTTGCTCTTGC-3'**
PotClf	**5'-GIVVIGTIGGIWSIGGIAARTCIAC-3'**
PotClr	**5'-ACICCRTTYTCDATDATRTTIG-3'**
PotHelr	**5'-GAICCRWA IGARTCIAIIACRTG-3'**
PotHelf	**5'-TGYGAYAAYCARYTIGAYIIIAAYG-3'**
Helf	**5'-GAATCTCTTTGCGTGATATC-3'**
Helr	**5'-CAGATGTGCAAAGTGCTGAA-3'**
LuHCr	**5'-GCCGCAACTGATCCACACTG-3'**
Lup700r	**5'-TGCAGTATTCTTCCACTCACA-3'**
Lup80r	**5'-CAGATGTGCAAAGTGCTGAA-3'**
Oligo (dT) 16	**5′-ACTATCTAGAGCGGCCGCTTT16–3'**

Primers used in screening of E.coli transformants and clones sequencing

M13F	**5'-GTAAAACGACGGCCAG-3'**
M13R	**5'-CAGGAAACAGCTATGAC-3'**
pJET	**5'-CGACTCACTATAGGGAGAGCGGC-3'**
pJET	**5'- AAGAACTACGATTTTCCATGGCAG-3'**

Primers used for amplifying plasmids in phytoplasma-infected samples

136	**5- AATAAACCCAACCTAAAACTGA-3'**
224	**5– TCCGTCGGGGTTTATTTCC-3'**
225	**5– TGCAGTTGTAATTGGTTGTC-3'**
226	**5—AAGATAAAACTCAATTCATTCCATGTGT-3'**
253	**5– AAGCAATAAAGGAATCTAATCTAATAAAATG3'**
254	**5– CACTCTTTTCTTTAATTTAATCTTCAT-3'**
423K8	**5 – TATATTTAAGATTTAATTATGC-3'**
423K9	**5– ACGTAGGTCATCTAAAATAATAC-3'**
426C7	**5– CTTCAGTATTAAACCATTGAG-3'**
426C8	**5–TCAAGAATGTATTATTTTAGATGACCTACG -3'**
503T0	**5– TACTTTAGTTGAGGTTTTATTTTCGCC-3´**

1.5.3 Equipment

ABI Prism 310 sequencer (PE Applied Biosystems, Foster City, CA, USA), iCycler (Bio- Rad), C1000 TM Termal Cycler (Bio-Rad), Swift maxi Cycler (ESCO), centrifuge MICRO 200R (Hittech, Germany), complect for electrophoresis Minis-150 (Biotech, Czech R.), digital dry bath AccubBlock TM (Labnet International, Inc.), thermostat BT120 (Lab System).

1.5.4 Purification of Virus-Like Particles

Virus-Like-Particles (VLPs) were purified from *Lupinus polyphyllus* showing mild mosaic symptoms as described previously in the 'Techniques in plant virology' protocol with some modifications. For extraction, after 18- 20 days of inoculation 100g of plants material was harvested and briefly, homogenized with 2 volumes (w/v) of 0.1M phosphate buffer pH 8.0 and 0.5% 2- mercaptoethanol. Then, filtered through 3 layers of gauze adding 10% (v/v) chloroform and 8% (v/v) 1-butanol mix and rotate for 30 min-1 hour, following centrifugation at 10000g for 10 min at 15 °C. The pellet was discarded, supernatant collected, and 4% NaCl and 4% PEG were added, mixed and kept under at 5-6 °C overnight and then, centrifuged at 10000g for 10 min. After processing a supernatant was removed and pellet resuspended in 0.01M phosphate buffer pH 7.5 (1/10 of the original volume) and optionally, following centrifugation at 3500g for 2 min or after removing a pellet, supernatant could be processing on sucrose cushions 20% (w/v) 1 ml per tube, and 160,400g for 90 min.

The pellet was resuspended overnight in 0.01M phosphate buffer pH 7.5 (1/50 orig. vol. and centrifuged at 7800g for 10 min and was processing with supernatant which was centrifuged in density gradients of sucrose (10 to 40%, w/v) in a swinging bucket rotor 96,500 for 2 hours. By fractionation, the virus band was collected in the gradient. Process fractions were done by dilutions 1:3 in 0.01M phosphate buffer pH 7.5 and sediment at 102,600g for 1 hour. The supernatant was removed and pellet was processed and resuspended in overnight in 0.01M phosphate buffer pH 7.5 1/100 of the original volume.

1.5.5 Transmission Electron Microscope (TEM)

A copper grid was placed on a drop of purified virus particles on a hydrophobic surface (a piece of parafilm), incubated at RT for 5-10 min, then washed with 40 drops of water, and dried with a piece of filter paper. The virus particles were stained with 2% (w/v) of uranium acetate for 2 min, dried as above, and observed using a transmission electron microscope.

1.5.6. Phenol extraction method of total nucleic acids

About 100-200 mg plant material was ground in liquid nitrogen using a mortar and pestle. To the ground tissue, 0.5 ml of 1x STE buffer containing was added, then equal vol. of phenol: chloroform: isoamyl alcohol (25:24:1) was added and tubes were incubated at room temperature for 5-6 min with centrifuged 14000g. Aquatic faze was transferred to a new tube and re-extracted with phenol: chloroform: isoamyl alcohol as described above if it was needed. To remove rest of phenol aquatic fraction containing nucleic acids was mixed with equal vol. of chloroform and centrifugated at 14000g for 2 min. Then nucleic acids were subjected for RNase or DNase digestion and then precipitated by equal vol. of isopropanol/ ethanol for 30 min at -70° C with adding 1/10 vol. of 3M N-acetate to acquired 0,3M N-acetate in final volume, then centrifuged at 14000g for 15 min. The pellet was washed with 0.5 ml 70% ethanol, dried for 5 min at 50 °C, and dissolved in 20-50 µl ddH2O.

1.5.7 Extraction of nucleic acids by using NucleoSpin® Extract II

RNA was isolated from infected leaves of *L. polyphyllus* with mild mosaic symptoms using a NucleoSpin RNA Plant Kit (Macherey-Nagel, Germany) according to the manufacturer's instructions as well as DNA from phytoplasma-infected samples.

1.5.8 cDNA production

RNA was isolated from infected leaves of *L. polyphyllus* with mild mosaic symptoms using a NucleoSpin RNA Plant Kit (Macherey-Nagel, Germany) or phenol/chloroform method according to the manufacturer's instructions and used as a template for cDNA synthesis. First-strand cDNA was synthesized either with an

iScript ™ cDNA synthesis kit (Bio-Rad) with random primers or with Monster Script™ cDNA synthesis (Epicentre, Technologies, USA) and an oligo (dT) primer according to the manufacturer's instructions. The samples were treated with RNase H for 15 min at 37 °C after transcription.

1.5.9 cDNA with specific primers

Specific primers were designed based on sequences' information derived from the alignment of known and completed potyviral sequences available in the GenBank for the genus *Potyvirus* (Tab 2.). The first strand synthesis following by PCR was performed with total nucleic acids extracted from infected material of *Lupinus polyphyllus* using various combinations of reverse primer sets. In addition, these specific primer sets were used to either extend or generate more DNA fragments with or without combinations of random primer.

1.5.10 PCR

PCR was done using DreamTaq polymerase (Fermentas, Lithuania) with 10X DreamTaq buffer. The reaction volume of 20-15 µ contained 0.5-1µ of cDNA, 0.5 µ of each primer each 20mM, 1 µ dNTPs 25mM, and 2.5 U of the enzyme. The PCR products were either sequenced directly with the corresponding primers or cloned into the pJET vector (Fermentas, Lithuania) according to the manufacturer's recommendations. Sequencing reactions with pJET forward and reverse primers were performed using a Big-Dye Terminator ver. 3.1 sequencing kit (Applied Biosystems, UK) and analyzed in an ABI Prism 310 sequencer (PE Applied Biosystems, Foster City, CA, USA). Contigs were assembled using DNA STAR software package (v.8.0.2) (Lasergene, Madison, WI, USA).The phylogenetic relationships of LP were analyzed using MEGA software (v. 4.1), (Tamura et al., and 2007).

Cycling conditions: an initial denaturation step at 94 ºC for 2 min, then 30 cycles at 94 ºC for 30 sec., X ºC for 30-50 sec (depends on primer annealing temperature), and 72 ºC for 1-3 min. A final extension step was performed at 72 ºC for 5 min.

1.5.11 Phi29 amplification or circular plasmid DNA

Whole-genome amplification using the unique property of the enzyme Phi29 DNA polymerase was used to generate copies of the genome. The enzyme was supplied with buffer of the following composition: 330mM Tris-acetate (pH 7.9 at 37 °C), 100mM Mg-acetate, and 660mM K-acetate, 1% (v/v) Tween 20, 10mM DTT; Reactions were done by using 10x reaction buffer in 20µl by adding oligo hexamer primer or specific one.

1.5.12 5´ RACE System for Rapid Amplification of cDNA ends

5´ RACE System Rapid Amplification kit was used according to manufacturer's instructions: ends (RACE) is used for the amplification of nucleic acid sequences from a template between a defined internal site and unknown sequences at either the 3' or the 5' -end of the mRNA.

In this work, first strand cDNA was synthesized from total or poly (A) + RNA using a gene-specific primer (GSP1) and iScript or Monster Script™ reverse transcriptase with reduced RNase H activity. After first strand cDNA synthesis, the original mRNA template was removed by treatment with the RNase Mix (mixture of RNase H). Then it was separated from cDNA using PCR clean up kit. A homopolymeric tail was then added to the 3'-end of the cDNA using TdT and dCTP. PCR amplification was accomplished using Taq DNA polymerase, a nested, gene-specific primer (GSP2, designed by us) that anneals to a site located within the cDNA molecule, and a novel deoxyinosine-containing anchor primer provided with the system.

5' RACE products were cloned into pJET or TOPO vector then amplified and sequenced. The Abridged Anchor Primer (AAP), Abridged Universal Amplification Primer (AUAP), Anchor Primer (AP) was supplied by kit (5´ RACE System).

1.5.13 Cloning and sequencing into pJET1.2 and TOPO vectors

PCR products were purified from the agarose gel or directly from the PCR tube with the NucleoSpin® Extract II and were ligated into the cloning vector pJET (Fermentas) or into TOPO vectors at a molar ratio of about 4:1 (PCR product: vector) according to manufacturer's instructions. Ligated plasmids were transformed into E. *coli* competent cells by means of heat shock at 42 °C for 30 sec after the heat, the tubes were incubated on ice for 2 min, then 200 µl of SOC medium were added,

and the cells were cultured for 1 h at 37 °C. The bacteria were cultured overnight on LB agar at 37 °C. White colonies were screened by PCR, using vector-based primers flanking the cloning site. Colonies with positive PCR results were cultured in LB-amp overnight at 37 °C. The experiment was repeated if it was needed. The sequences were determined using the Sanger sequencing with an ABI 3730XL sequencer (Eurofins MWG Operon) and assembled into contigs using the DNA STAR software package (v.8.0.2) (Lasergene, Madison, WI, USA).

1.5.14 DNA sequencing

PCR products and/or its clones containing inserts were selected and subjected for bidirectional sequencing using pJET, M13 reverse and forward primers, and gene specific primers using a Big-Dye Terminator ver. 3.1. sequencing kit (Applied Biosystems, UK) and analyzed in an ABI Prism 310 sequencer (PE Applied Biosystems, Foster City, CA, USA) Sequences of DNA obtained were then compared to known plant virus groups using the BLAST program (http://www.ncbi.nlm.nih.noBLAST/) and also used to align partial DNA fragments with each other by using DNASTAR software package (v.8.0.2) (Lasergene, Madison, WI, USA).

1.5.15 Data analysis and accession numbers

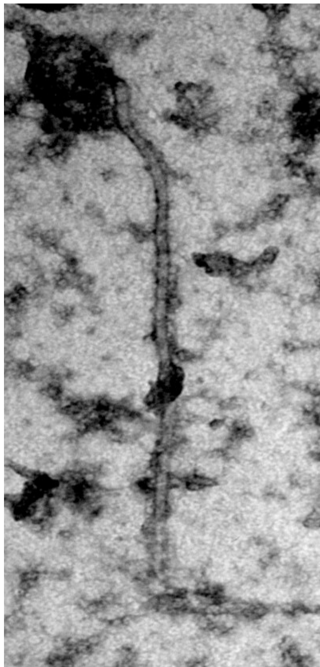
Sequence similarity searches were performed in the National Center for Biotechnology Information (NCBI) using the BLAST program. Multiple alignments of nucleic and amino acid sequences were carried out using the DNASTAR software package (v.8.0.2) (Lasergene, Madison, WI, USA) with the default parameters. Phylogenetic trees were constructed with the Molecular Evolutionary Genetic Analysis MEGA4 program (Tamura et al*.,* 2007). A bootstrap test was conducted with 100 replicates for the neighbor-joining (NJ) tree. Viral sequences, other than LuMV were retrieved from the GenBank (Tab. 1). Primers were designed using the DNASTAR software package (v.8.0.2), (Lasergene, Madison, WI, USA).

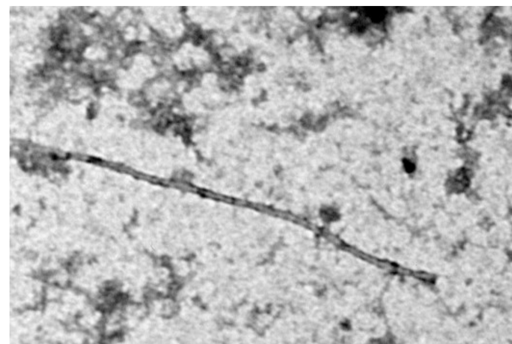
Chapter 2
Results

2.1. Molecular analysis of *Lupine mosaic virus*, LuMV – a new member in genus *Potyvirus*

Our aim was the screening samples collected in nature from wild and cultured plants, predominately of family *Fabaceae* showing mosaic, yellowing, and chlorotic, stunting symptoms. For the detection the reverse transcription and PCR were used with universal and degenerate specific primers for testing viruses in following genera: *Potexvirus, Potyvirus*, *Comovirus, Carlavirus* and *Cryptovirus* (data not shown). These primers have been designed by us or known specific primers have been used as well, f.i. for potyvirus detection Gibbs, (1977). Testing samples showing symptoms likely with viral infection were subjected for the detail analysis and among them - the garden lupine with mild mosaic symptoms and yellowing was assumed to have a viral infection (Fig.1a). For the amplification specific primers (Gibbs et al., 1977) were used and 2.7 kbp PCR product was generated (Fig.6), then sequenced and confirmed the potyvirus infection. Thus, a *Lupine mosaic virus*, LuMV was detected and taken for the complete sequencing because of primary promising results that proposed the presence of a new putative member in genus *Potyvirus*. This sample was subjected to the electron microscopy examination by which was confirmed a presence of filamentous virus particles of 690 nm in length which were visible in negative-contrast stained plant sap preparations (Fig. 1b.), (Sarkisova et al., 2009). It was in correspondence to potyviruses which are known as flexuous, rod or filamentous shaped particle compose of a monopartite single-stranded positive RNA about 9-11kb long encapsidated by approximately 2000 copies of coat protein (CP) (Matthews, 1991).The newly isolated virus was maintained by a mechanical passage in *Nicotiana benthamiana* and *Chenopodium quinoa* causing mild mosaic symptoms on non-inoculated leaves.

Fig.1a Specimen of the plant *Lupinus polyphyllus* with mosaic symptoms





1b) Preparation from plant sap in electron microscope with flexuous elongated particles 690 nm long

The complete nucleotide sequence of the LuMV ssRNA was determined by amplifying and cloning of partial segments of the virus genome; reverse transcription and PCR, using degenerate and/or specific oligonucleotide primers has been done. It was shown that LuMV genomic RNA consist of 10,113 nt in length and contains one open reading frame (ORF) encoding a polyprotein of 3,213 amino acids with a molecular mass (Mr) of 364.3 kDa. The complete nucleotide sequence was submitted to GenBank under accession number HM748648 and analysis of the deduced aa

sequence was performed by comparison to other available potyvirus sequences (Sarkisova et al., 2011).

2.2. Analysis of LuMV RNA sequence

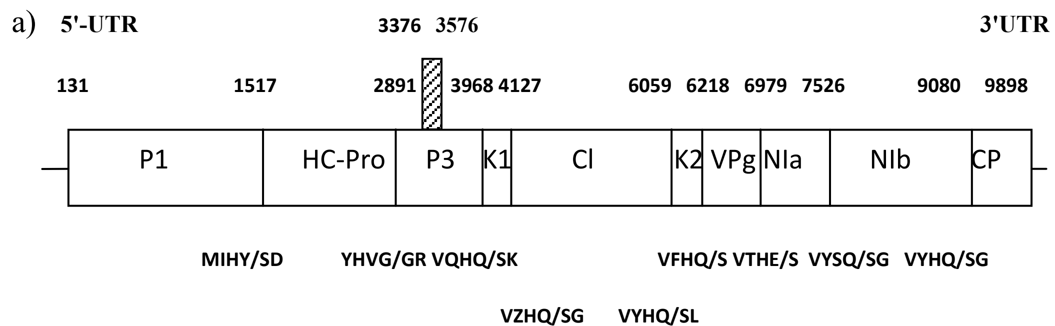
The complete sequence of *Lupine mosaic virus*, LuMV was taken by a PCR based approach using specific (Gibbs et al., 2003), and degenerate potyvirus primers either from cloned PCR products or directly which were designed and derived from conserved motives of other potyviruses. The first strand synthesis and PCR reactions with random and specific primers generated many fragments ranging in size from 200 bp to 2 kb (Fig.2b); they were aligned with each other based on their overlapping sequence (Fig. 2b). The RNA is 10,113 nt long excluding 3'- terminal poly (A) tract. The overall composition of CG of LuMV ssRNA was 39.9 % (Sarkisova et al., 2011). A *Lupine mosaic virus*, LuMV was very similar in size and genome composition to that of other potyviruses (Puurand et al, 1994). The sequence data were used to align overlapping partial DNA fragments and to compare levels of sequence homology with have already known viruses of genus *Potyvirus* by using BLAST and the DNASTAR software package (v.8.0.2), (Lasergene, Madison, WI, USA) for creating multiply alignments in order to see % of identities following by phylogenetic analysis to know their relatedness.

The analysis of LuMV complete sequence was done and showed four in frame initiation codons within the first 300 nt of the long single open reading frame (ORF) at nt positions 131–133, 254–256, 260–262, and 293–295. However, only third triplet was in a favorable context (GACAATGGC), which were taken by us in our lupine potyvirus, LuMV, as closest to the consensus sequence AACAAUGGC for translational initiation in plants (Lutcke et al., 1987) and thus most likely the initiator of LuMV polyprotein in vivo. The UAA terminative codon was located at 9899–9901nt and the 3'- untranslated region consists of 212nt (Sarkisova et al., 2011). The recently described putative protein PIPO (Chung et al., 2008) which is produced by a translational frame shift near the middle of the P3 cistron was identified by using software package MLOGD in the +2 reading frame starting from a GAAA motive at position 3376 (Sarkisova et al., 2011). This motive is different from that of the highly conserved (G) GAA AAA A (A) motive which is generally known in this frame for other potyviruses (Chung et al., 2008) and ends with a UAA termination codon at

position 3574–3576. The predicted protein is 66 aa long with a (Mr.) of 7.3 kDa (Fig.2a) (Sarkisova et al., 2011). It forming a complex with the CI protein that proposed to play a role in the interaction with the plasmodesmata and thus, providing cell-to-cell virus movement, reading frame starting from a GAAAA motive (Chung et al., 2008; Wei et al., 2010b; Wen et al., 2010). The *Turnip mosaic virus*, TuMV was the first potyvirus where PIPO protein was identified and later was shown it to be presented in all known potyviruses available in GenBank at that moment. Up to present, this protein has been found in other genera of family *Potyviridae* (Chung et al., 2008).

If we looking at the genetic diversity of distinct potyvirus species by comparison their complete genome sequences is visible that the first protein (P1), the third protein (P3) and the N-terminal domain of the CP are the most variable regions of the potyviral polyprotein, while the replicase (nuclear inclusion b, NIb) is the most conserved protein (Shukla et al., 1991).

Lettuce mosaic virus, LMV, *Plum pox virus*, PPV and *Panax virus Y*, PanVY were found to be the closest relatives among the 63(number of available data at that time) and based on the data of the phylogenetic analysis of available potyviral sequences shared polyprotein aa sequence identities 54, 53 and 52% to LuMV polyprotein, respectively (Sarkisova et al., 2011). Fortunately, the existing criteria for the demarcation of potyviruses at the species level did allow for precisely establishing the taxonomic relationship between *Lupine mosaic virus*, LuMV and other completely sequenced potyviruses. The demarcation criteria was less than 80% of aa sequences for coat protein (CP) and 85% for the whole genome what pointed for the new potyvirus species, according to ICTV for genus *Potyvirus* (Adams et al., 2005; Desselberger et al., 2009).



b)

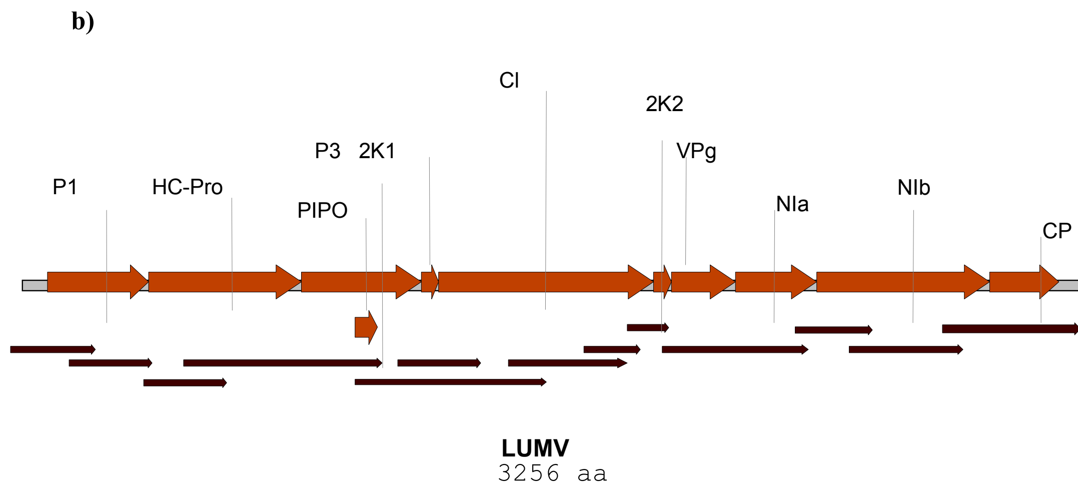


Fig. 2 Genome organization of *Lupine mosaic virus* (LuMV) The genomic RNA (10.113 kb) encodes a large polyprotein that is proteolytically processed into P1, helper component proteinase (HC-Pro), P3, cylindrical inclusion (CI) protein, 6K (6 kDa), viral genome-linked protein (VPg), nuclear inclusion proteins (NIa and NIb), and coat protein (CP) (Adams et al., 2005; Ng and Falk, 2006). a) Genomic RNA b) schematic representation of the viral genomic RNA with overlapping amplified segments

2.2.1 Proteolytic processing of the polyprotein

It is known that mature potyviral proteins are produced via proteolytic processing of the viral precursor polyprotein by tree virus encoded proteases. Cleavage sites in the C-terminal part of this polyprotein on 2/3 are processed by NIa protease whereas the P1 and HC proteins are auto catalytically release from the polyprotein precursor (Dougherty et al., 1989; Carrington et al., 1989a; Verchot et al., 1991). By using comparison of both determined and putative consensus sites in known polyproteins seven putative cleavage sites were identified in *Lupine mosaic virus*, LuMV polyprotein sequence (Fig.2a) and these sites occur at the expected positions what gave us to predict size of cleavage proteins based on data upon other potyviruses (Sarkisova et al., 2011). The serine protease catalytic residues H_{370}-D_{379}-S_{413}-G_{438} have been perfectly conserved in the C-terminal part of the P1 protein. The cleavage site between the P1 and the HC-Pro protein was found to be located on 24 aa downstream from the conserved G_{438} in the MIHY/SD site. The resulting protein consisting of 419 aa in length and shared about 26% aa sequence identity with that of *Pea seed-borne mosaic virus*, PSbMV. The HC-Pro/P3 cleavage site (G/G) was

found in the YHVG920/G921 context (Sarkisova et al., 2011). The consensus cleavage motive of the NIa-Pro protease responsible for proteolytic processing of the polyprotein precursor at seven sites had been predicted as DxVxHQ/S. This motive is not unique for LuMV and presented as well in the related *Lettuce mosaic virus* (LMV). Variation in the motive 6K2/VPg, where the cleavage site is Q/G, and in VPg/NIa, where the cleavage site is E/S occurred in our new potyvirus (Sarkisova et al., 2011).

2.2.2. Capsid protein (CP)

Amongst all genes of the potyviral, genome the CP is the most frequently selected for studies of genetic diversity and constructs of phylogenetic trees to clarify taxonomy. RT-PCR was done using potyvirus-specific primers. The sequence comparison through BLAST revealed the presence of a new species in genus *Potyvirus*. The CP coding region was 819 nt long terminated with UAA and coded for a protein of 273 amino acids (Sarkisova et al., 2011). BLAST search of this sequence fragment reveal similarity to the CP of viruses in the family *Potyviridae*, genus *Potyvirus* (Tab.2). Primary, using 2.7 kbp product was generated which spanned not only capsid protein (CP) but also 3'- untranslated region and part of *rep* (replicase) region by RT-PCR (Fig.7) (Sarkisova et al., 2009). BLAST showed homology at the aa (amino acid level) to *Narcissus yellow stripe virus,* NYSV, *Pepper veinal mottle virus,* PVMV, and *Chilli veinal mottle virus*, ChiVMV and so they were identified as closely related viruses with 64.4%, 64.3%, and 63.1% amino acid identity, respectively(Sarkisova et al., 2009). The multialignment comparison coat proteins (CP) coding regions of available 63 potyvirus sequences was done and phylogenetic analysis in order to estimate phylogenetic relations more surely. The highest aa entire amino acids identity was 64 % to the CP region of *Lettuce mosaic virus,* LMV and identities ranged from 60 % to 64% to CP regions of potyviruses and with *75* to 79% between conservative aa inside of the capsid protein (Tab.2). The highest % of identities in nt level was shown to the *Lettuce mosaic virus,* LMV either between other close related viruses was in range no more than 64% (Tab.3).Thus, based on demarcation criteria for conservative CP region, which is under 80%, the result indicated a new virus (Adams et al., 2005; Berger et al., 2005; Desselberger et al., 2009). The phylogenetic tree (Fig.4) and % of identities comparison of the closest related viruses is shown in (Tab. 3). The conservative motive DAG has been found in

N-terminal part of CP at 12-14aa that indicated on aphid transmission (Fig.8) (Sarkisova et al., 2011).

Tab. 2 Sequence identity comparison (%) of entire amino acids (above the diagonal) and conserved only (bellow the diagonal) of CP between LuMV and other close related viruses

	LuMV	LMV	PTV	PVY	PVV	VVY	WPMV	SuCMoV	PepSMV	PepMoV
LuMV		**64**	**61**	**63**	**60**	**62**	**60**	**61**	**63**	**63**
LMV	**79**		**62**	**64**	**62**	**64**	**62**	**64**	**68**	**62**
PTV	**76**	**75**		**74**	**89**	**73**	**87**	**74**	**75**	**74**
PVY	**78**	**78**	**84**		**70**	**77**	**71**	**79**	**81**	**74**
PVV	**75**	**74**	**95**	**83**		**71**	**85**	**72**	**72**	**71**
VVY	**76**	**75**	**83**	**86**	**82**		**70**	**78**	**76**	**75**
WPMV	**75**	**75**	**95**	**84**	**91**	**83**		**74**	**73**	**75**
SuCMoV	**77**	**78**	**86**	**90**	**84**	**87**	**85**		**79**	**77**
PepSMV	**77**	**81**	**85**	**88**	**83**	**87**	**86**	**88**		**74**
PepMoV	**77**	**77**	**86**	**83**	**85**	**84**	**86**	**85**	**84**	

***full viral names including abbreviations in Tab.1**

Tab. 3 Nucleotide sequence identity comparison (%) of CP between LuMV and other close related viruses

	LuMV	LMV	PTV	PVY	PVV	VVY	WPMV	SuCMoV	PepSMV	PepMoV
LuMV		**63**	**61**	**62**	**63**	**62**	**62**	**60**	**61**	**63**
LMV			**61**	**62**	**62**	**66**	**62**	**63**	**66**	**60**
PTV				**70**	**81**	**69**	**81**	**69**	**70**	**70**
PVY					**67**	**71**	**69**	**76**	**73**	**68**
PVV						**69**	**79**	**68**	**70**	**70**
VVY							**70**	**70**	**73**	**70**
WPMV								**69**	**70**	**73**
SuCMoV									**73**	**70**
PepSMV										**71**
PepMoV										

***full viral names including abbreviations in Tab.1**

2.2.3 Replicase (*rep*)

The NIb is known as the RNA-dependent RNA polymerase or replicase, which is known to be responsible, and involved in the virus replication (Domier et al., 1987). A homology search with the deduced amino acid sequence showed a significant homology to the replicase regions of other potyviruses and all nine motives (Ia to VIII) of the RNA-dependent RNA polymerases (Koonin et al., 1991). It has been presented in the NIb protein of *Lupine mosaic virus*, LuMV (Fig.3), but with slight variation with the sequence SLMV2788 instead of TLMV in motive V. The highest amino acid sequence homology of the replicase region was to the *rep* of *Pea seed-borne mosaic virus*, PSbMV, with 59% identities and 78% to *Pennisetum mosaic virus*, PenMV for conserved amino acids only inside replicase region (Tab.4) (Sarkisova et al., 2011). Sequence comparisons for the entire *rep* motive at the nt level revealed high homologies of *Lupine mosaic virus*, LuMV to *Freesia mosaic virus*, FreMV, *Pea seed-borne mosaic virus,* PSbMV, *Cocksfoot streak virus*, CSV, *Sorghum mosaic virus*, SrMV, *Maize dwarf mosaic virus*, MDMV, *Pennisetum mosaic virus*, PenMV and *Sugarcane mosaic virus*, SCMV in the range from 42 to 63% (Tab.5). The highest aa% of identities was 59 to *Pea seed-borne mosaic virus,* PSbMV. Both nucleotide and amino acid sequences comparison of the entire

replicase motive including the % of identities of entire and only conserved in the replicase of *Lupine mosaic virus*, LuMV and other closely related viruses are illustrated in Tab. 4 and 5. All calculations were done by using the DNASTAR software package (v.8.0.2), (Lasergene, Madison, WI, USA). The closest related for comparison have been chosen based on phylogenetic comparison of rep regions of 63 potyviral sequences available in GenBank.

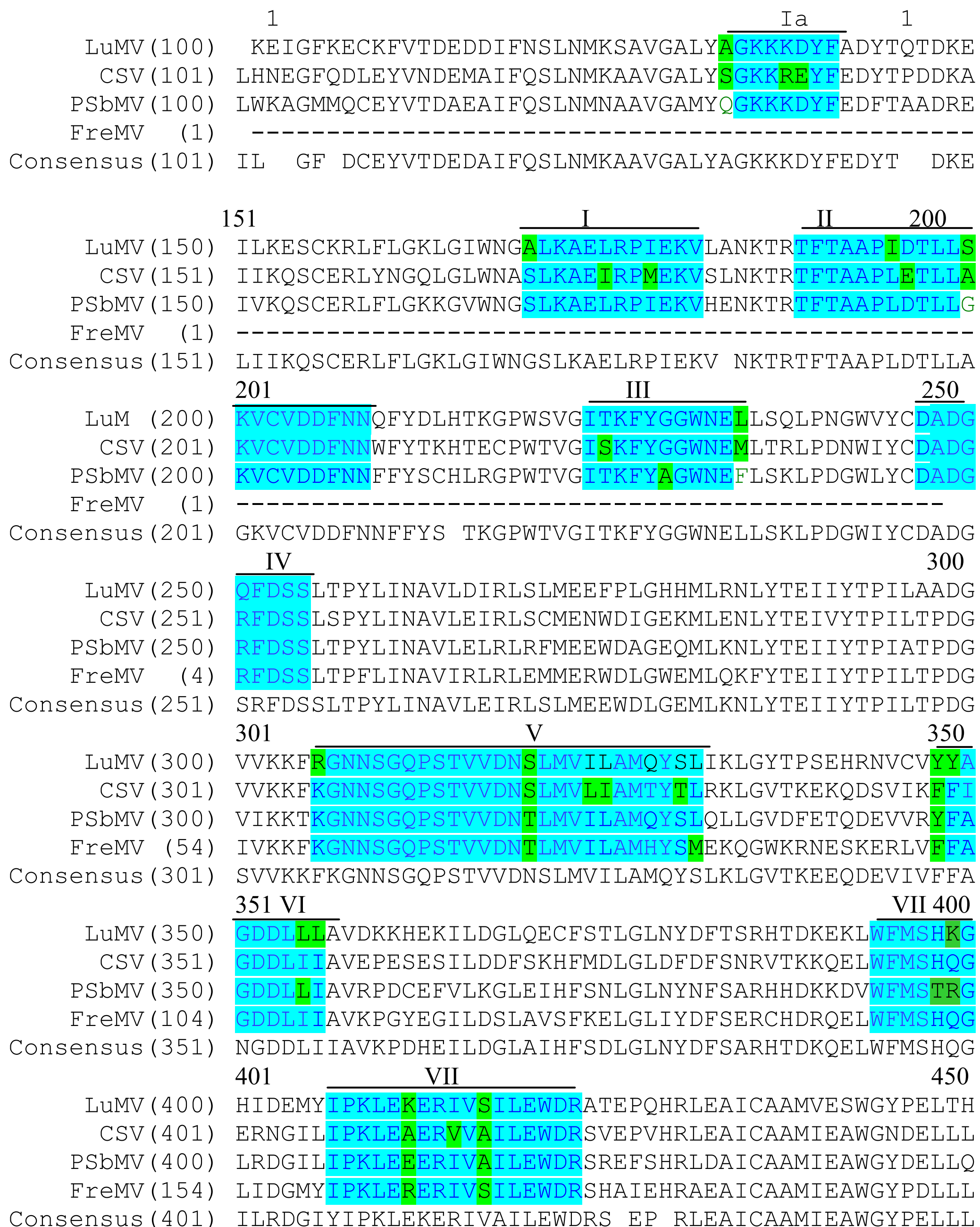


Fig.3. Comparison of conserved motives (Ia -VIII) of the rep of LuMV to that of other close related *viruses* (partial sequence with conservative motives*).* The following viruses were included in the analysis: cocksfoot streak virus (CSV), pea seed-borne mosaic virus (PSbMV), freesia mosaic virus (FreMV)

Tab. 4 Sequence identity comparison (%) of entire amino acids (above the diagonal) and conserved only (bellow the diagonal) of *rep* between LuMV and other close related viruses

	LuMV	FreMV	PSbMV	CSV	SrMV	MDMV	PenMV	SCMV
LuMV		55	59	48	55	53	55	53
FreMV	72		53	51	55	57	56	56
PSBMV	73	72		49	56	53	55	55
CSV	76	69	66		50	48	48	49
SrMV	74	72	73	67		81	76	70
MDMV	64	71	74	66	90		76	74
PenMV	78	72	73	66	86	87		72
SCMV	74	68	69	65	79	83	81	

*full viral names including abbreviations in Tab.1

Tab. 5 Nucleotide sequence identity comparison (%) of *rep* between LuMV and other close related viruses

	LuMV	FreMV	PSbMV	CSV	SrMV	MDMV	PenMV	SCMV
LuMV		42	62	60	63	63	63	61
FreMV			42	40	43	42	44	43
PSBMV				60	63	63	63	61
CSV					61	61	63	62
SrMV						72	72	71
MDMV							71	70
PenMV								72
SCMV								

*full viral names including abbreviations in Tab.1

2.2.3.1 Phylogenetic relationships

Genetic relatedness was analyzed by comparing amino acid sequences of the conserved motives of capsid protein (CP) and replicase (*rep*) of *Lupine mosaic virus,* LuMV to known viruses. Based on multiple alignments of these regions, a phylogenetic trees were drawn (Fig. 4, 5) by using MEGA software (v. 4.1) (Tamura et al., 2007) all abbreviations in (Tab. 11). A closer relationship was shown between LuMV CP to *Lettuce mosaic virus*, LMV, *Peru tomato virus*, PTV, *Potato virus Y*, PVY, *Potato virus V*, PVV, *Wild potato mosaic virus*, WPMV, *Sunflower chlorotic mottle virus*, SuCMoV, *Pepper severe mosaic virus* , PepSMV, *Verbena virus Y*, VVY and *Pepper mottle virus*, PepMoV (Fig. 5); all abbreviations in (Tab 11). In case, of *rep* motive comparison the phylogenetic analysis revealed that the closest potyvirus is *Freesia mosaic virus*, FreMV and other viruses, which belong to one cluster sharing the same root are illustrated by (Fig.6).

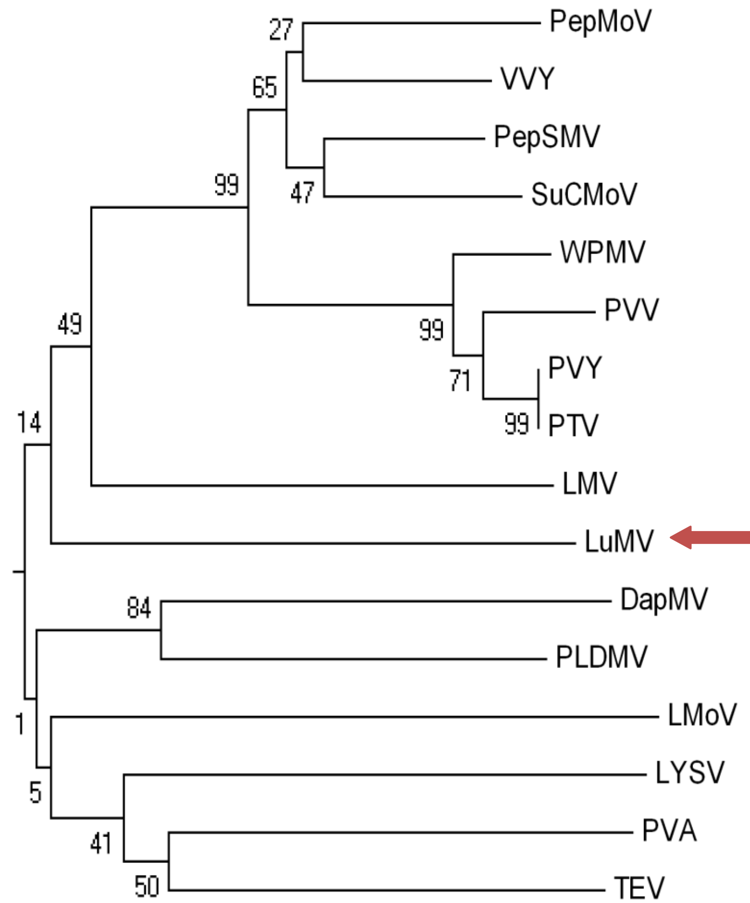


Fig.4 Phylogenetic relationships among the CP of LuMV and other close related potyviruses. Analysis was done and bootstrapped neighbor-joining tree was drowning with using Mega v. 4.1 (Tamura K et al., 2007) program. Bootstrap values are shown on the branches of the tree. Accession numbers of CPs are indicated in brackets in the legend. Virus notations are: Japanese yam mosaic virus (JYMV; NP_734232), konjac mosaic virus (KoMV; YP_529499), lettuce mosaic virus (LMV; NP_734162), narcissus yellow stripe virus (NYSV; YP_002308564), panax virus Y (PanVY; YP_003725728), plum pox virus (PPV;V), scallion mosaic virus (ScMV; NP_734132), sweet potato feathery mottle virus (SPFMV; NP_734318), turnip mosaic virus (TuMV; NP_734222), potato virus A (PVA; NP_734368), yam mosaic virus (YMV; YP_022760), tobacco etch virus (TEV; NP_734206)

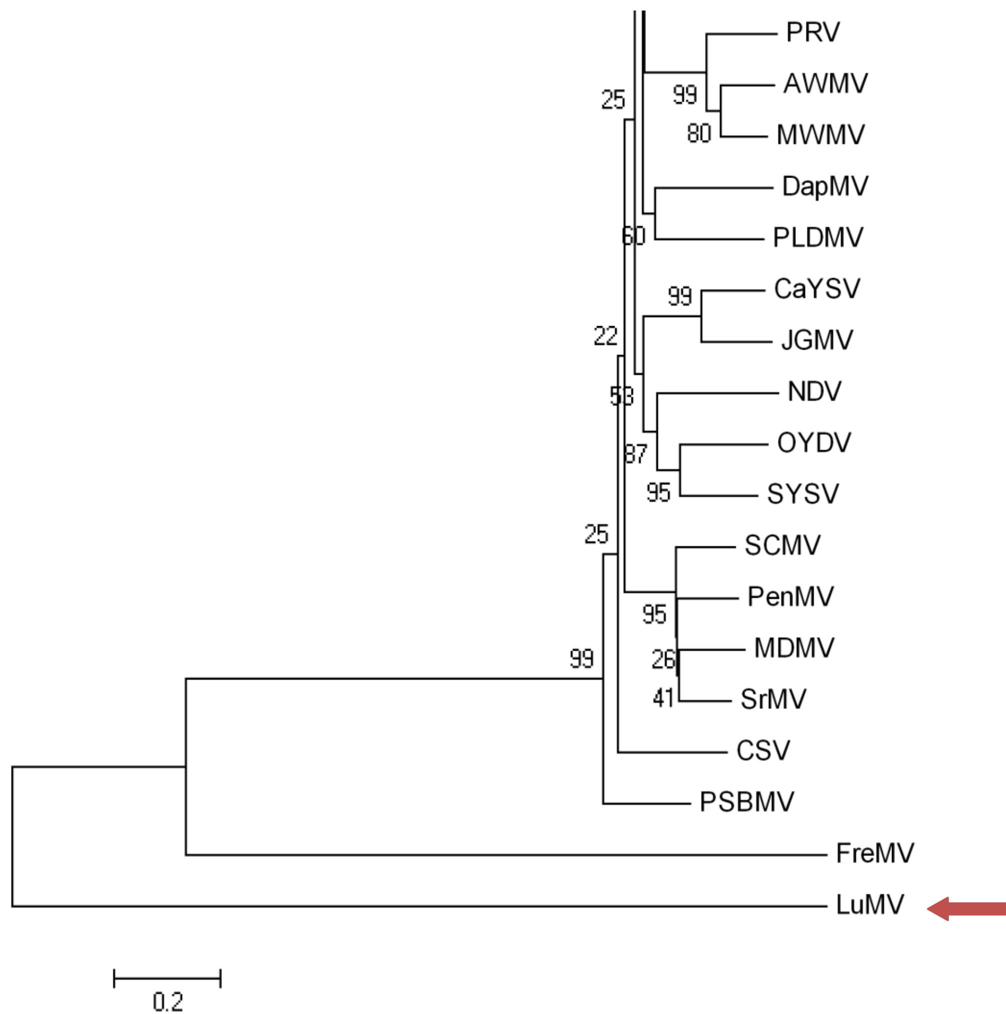


Fig.5 Phylogenetic relationships among *reps* of LuMV and other close related potyviruses. Analysis was done and neighbor-joining tree was drowning with using Mega 4.1 (Tamura et al, 2007) program. Bootstrap values are shown on the branches of the tree. Accession numbers of *reps* are indicated in brackets in the legend. The scale bar represents a genetic distance of 0.02. Virus notations are: freesia mosaic virus (FreMV; YP_003620391.1), pea seed-borne mosaic virus; NP_734427), cocksfoot streak virus (CSV; P_734397), sorghum mosaic virus (SrMV; P_734091), maize dwarf mosaic virus (MDMV; NP_734151), pennisetum mosaic virus (PenMV; NC_007147), sugarcane mosaic virus (SCMV; NP_734141), shallot yellow stripe virus (SYSV; P_331422), onion yellow dwarf virus (OYDV; NP_871746), narcissus degeneration virus (NDV; YP_001019196), johnsongrass mosaic virus (JGMV; NC_003606), canna yellow streak virus (CaYSV; YP_003208059), Papaya leaf-distortion mosaic virus (PLDMV; NC_005028), daphne mosaic virus (DapMV; YP_611119), Moroccan watermelon mosaic virus (MWMV; YP_001552429), Algerian watermelon mosaic virus (AWMV; YP_001936195), papaya ringspot virus (PRSV; NP_734241),

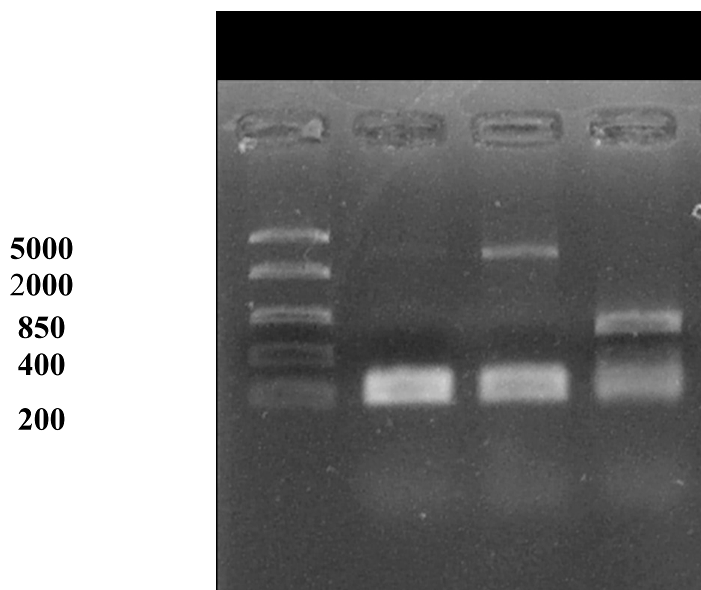


Fig.6. Gel electrophoresis of RT- PCR amplification products by using specific primer' sets for carla-, potex-, and potyviruses detection: (lane 1 - marker), (lane 2 -carlavirus), (lane 3- potexirus) and (lane 4- potyvirus), were successfully generated one RT-PCR fragment 2.7 kbp - the sequenced region spanned C-terminal part of the viral replicase, CP and 3'-UTR.

2.2.4 Role of helper component proteinase (HC-Pro) and capsid protein (CP) in potyviruses

HC-Pro and CP motives are both known by several conservative regions, which presented in each potyvirus, were recognized for *Lupine mosaic potyvirus*, LuMV. The role HC-Pro protein of potyviruses was known be related with systemic movement, aphid transmission, and polyprotein processing. In addition, the central and C-terminal regions of potyviral HC-Pro was supposed to be significant in suppression of posttranscriptional gene silencing (PTGS) as indirect result between long-distance movement and genome maintenance (Kasschau et al., 2001; Varrelmann et al., 2007). Silencing suppression was firstly studied on potyvirus TEV HC-pro (Varrelmann et al., 2007).

Multiple functional domains have been identified within 457 aa long LuMV HC-Pro protein (Fig.7) with molecular weight of 50.8 kDa; in the C-terminal part of this protein was observed a conservative motive, C_{641}-X_{72}-H_{71} (Sarkisova et al., 2011) which is known to be responsible for the protease activity of HC (Atreya et al., 1992). The CCCC amino acid motive in the central region of LuMV HC-Pro was found to be located at the positions 292-294 aa (Sarkisova et al., 2011) (Fig.7).

The LuMV IGN motive was located at the position 250-253aa within central region of the HC-Pro (Fig. 7) (Sarkisova et al., 2011). Its functions were shown to be associated with genome amplification in (Cronin et al., 1995) and as was shown on *Pea seed-born mosaic virus,* PSbMV, HC-Pro can be a major determinant of seed transmission (Johansen et al., 1996). In addition, for the sequence -specific interactions required motives KITC, PTK of HC-Pro and a single motive DAG_{12} in CP were presented in LuMV HC-Pro (Fig.7, 8). The N-terminal KITC motive was located at positions 53-56 aa in our LuMV HC-Pro (Fig.7) as was shown previously could mediate transmission and the sequence -specific interactions required of motive (Oh CS, et al., 1989; Blanc et al., 1998; Seo et al., 2010).

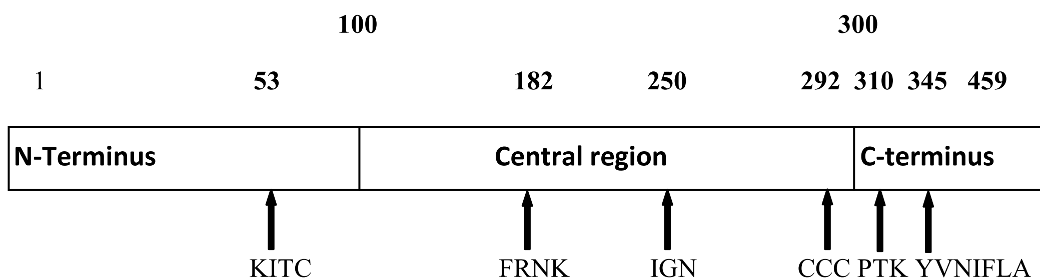


Fig.7 The helper component proteinase (HC-Pro) of LuMV has been divided into three functional domains based on similarities to HC-Pro of other potyviruses with the conserved amino acid motives (below) and the corresponding amino acid positions (above) indicated

Two motives FRNK, and PTK, which are also required for potyvirus transmission by aphids, were presented in the LuMV HC-Pro protein, which known to be associated with symptom expression as well (Varrelmann et al., 2007). These two components FRNK and PTK in LuMV were located at positions 182-185 aa and 310-312 aa (Sarkisova et al., 2011) (Fig.7). They have been discussed previously to be involved in aphid transmission (Huet et al., 1994; Gal-On, et al., 2000; Urcuqui-Inchima et al., 2001). The second motive PTK was discussed by authors (Blanc et al., 1998; Peng et al., 1998) binding with the N-terminal DAG motive of coat protein (CP), for LuMV CP which is located at the positions 12-14 aa in CP (Fig. 7, 8). In the N- terminal part of LuMV HC-Pro has been found a possible zinc finger binding motive $C\text{-}X_8\text{-}C\text{-}X_{13}\text{-}C\text{-}X_4\text{-}C\text{-}X_2\text{-}C$ (Aleman et al., 1996) starting at the position aa 27 to 58 as was shown for other potyviruses. All conservative motives were found in the LuMV HC-Pro as in other representatives of this genus.

HC-Pro and CP are needed for the aphid transmission in potyvirus group (Atreya et al., 1990; Ng and Falk, 2006; Varrelmann et al., 2007, Kasschau et al., 2001).

BLAST search following by phylogenetic analysis of LuMV HC-Pro amino acids sequences revealed similarity to *Potato virus A*, PVA, *Chili veinal mottle virus*, ChiVMoV, *Tobacco vain-banding mosaic virus*, TVBMV, *Pepper veinal mottle virus*, PVMV, *Tobacco etch virus,* TEV, and *Panax virus* Y, PanVY, *Yam mosaic virus*, YAM, *Lettuce mosaic virus*, LMV. They forming one cluster, the last three potyviruses showing closer relatedness by forming a subclade in the cluster (data not shown). The aa comparison showed the identities % to the corresponding protein of them and entire amino acid identities ranged from 51 % to 58% to HC-Pro region and with 69 to 75% between conservative inside of LuMV HC-Pro protein (Tab.6).

```
                1                                                 50

     LuMV (1)   SGTPKNDDKKLDAG---KNENKGKE--------ESTESSNKGAVIKDRDV
      LMV (1)   ------VDTKLDAGQGSKNDDKQKSSADSKDNVITEKGSGSGQVRKDDDI
 Consensus(1)          D KLDAG   KND K K         V S   S  G V KD DI
                51                                               100
    LuMV (41)   NAGTAGTHTVPRLKSITNKMRLPKTKKGVVLNIEHLLTYTPSQEDISNTR
     LMV (45)   NAGLHGKHTIPRTKAITQKMKLPMIRGKVALNLDHLLEYEPNQRDISNTR
   Consensus    NAG  G HTIPR KAITNKMKLP   K  V LNIDHLL Y P Q DISNTR
                101                                              150
    LuMV (91)   STQSQFNSWYENVKHDYDVQDDAMQIILNGFMVWCIENGTSPNISGVWTM
     LMV (95)   ATQKQYESWYDGVKNDYDVDDNGMQLILNGLMVWCIENGTSPNINGTWVM
Consensus       ATQ QF SWYD VK DYDV D AMQIILNG MVWCIENGTSPNI G W M
                151                                              200
    LuMV(141)   MDGEEQVEYPLKPMVEHAKPTLRQIMAHFSNAAEPYIEMRNTERSYMPRY
     LMV(141)   MDSEEQVEYALKPIIEHAKPTFRQIMAHFSDAAEAYIEMRNKKKPYMPRY
   Consensus  MD EEQVEY LKPIIEHAKPT RQIMAHFS AAE YIEMRN K YMPRY
                201                                            250
    LuMV(191)   GRQRNLNDRSLARFAFDFYEITSNTPEKAREAHFQMKAAALRDARNKMFG
     LMV(195)   GRLRGLNDMGLARYAFDFYETTSATPNRAREAHNQMKAAALVGTQNRLFG
   Consensus  GR R LND  LARFAFDFYE TS TP KAREAH QMKAAAL    NKLFG
                251                         284
   LuMV (241)   LDGKTGIQGEDTERHTAADVSADMHSLLGMRGT-
    LMV (245)   MDGGGSTQEENTERHTAADVNQNMHTLLGVRGLH
   Consensus     LDG     Q E TERHTAADV   MHSLLGMRG
```

Fig.8 Deduced amino acid sequences of the coat protein region of LuMV to the closest potyvirus - lettuce mosaic virus (LMV). In color shown conservatives motives inside CP and by red color marked -DAG

Tab.6 Sequence identity comparison (%) of entire amino acids (above the diagonal) and conserved only(bellow the diagonal) of HC-Pro between LuMV and other close related viruses

	LuMV	LMV	PVA	YMV	TEV	PanVY	SPFMV	PVMV	ChiVMV	TVBMV
LuMV		**56**	**57**	**54**	**51**	**58**	**56**	**57**	**58**	**58**
LMV	**73**		**52**	**54**	**51**	**58**	**54**	**55**	**52**	**53**
PVA	**73**	**70**		**53**	**52**	**55**	**57**	**53**	**56**	**56**
YMV	**73**	**72**	**72**		**49**	**53**	**52**	**52**	**52**	**52**
TEV	**69**	**68**	**70**	**63**		**49**	**52**	**53**	**53**	**54**
PanVY	**72**	**76**	**72**	**73**	**67**		**54**	**55**	**54**	**53**
SPFMV	**72**	**72**	**75**	**71**	**70**	**71**		**56**	**56**	**55**
PVMV	**75**	**71**	**73**	**72**	**68**	**72**	**72**		**79**	**71**
ChiVMV	**76**	**70**	**74**	**74**	**69**	**71**	**73**	**90**		**67**
TVBMV	**74**	**70**	**71**	**72**	**69**	**70**	**71**	**82**	**82**	

***full viral names including abbreviations in Tab.1**

2.2.5 P3 protein

Up to present, is not much known about P3, which is the third protein of the potyviral polyprotein precursor; analysis of the deduced aa sequence of LuMV P3 protein showed a high content of leucine which makes this 39.8 kDa protein (Sarkisova et al., 2011) very hydrophobic. Interestingly, despite the low level of conservation of this protein (Fig. 9) among potyviruses, some of leucine residues are organized in three conserved clusters located in the N-terminal region of potyviruses P3 proteins. These leucine-rich clusters were shown to present in tandems (Kobe et al., 1994). In case of LuMV P3, these are L_{994}-X_{13}-L_{957}-X_{6}-I (L) $_{965}$-X_{9}-L_{975}, L_{1005}-X_{2}-L_{1008}-X_{9}-L_{1018}, and L_{1053}-X_{12}-L_{1066}-X_{23}-L_{1090}-X_{6}-L_{1097}. In the first conservative motives, the second leucine residue is replaced by isoleucine (Fig.9). Leucine-rich repeats have been indentified in over sixty proteins; it might play a role in protein-protein interactions and in regulation of the viral transcription as was mentioned previously by Kobe (et al., 1994). Multiply alignment comparison revealed that *Papaya ringspot virus*, PRSV and *Plum pox virus*, PPV

as the closest related viruses with highest % of entire amino acids identities 31% and 51 to 52 % for conserved only inside P3 protein, respectively (data not shown).

```
               1                                                  50
     LuMV  (1) RTDNLDPSALGNGNNRTNIMSQLIRSIYRPNEMHRMLTEHPLYIVYAML
     PRSV  (1) ----------GEFDPTTNCLHQLIRVIYKPHELRSLLRNEPYLIVIALM
     PPV   (1) --------LEVDKCDEFKNVKLLIRSIYKPQIMEQVLKEEPYLLLMSVL
Consensus                GD    TN L QLIRSIYKPNEM LLKEEPYLIVIALL
               51                                                  100
    LuMV  (51) GVLIALFNSGSLDKMADMYVRSDMDLSAVASSLTLLAHKVSIAKTLEF
    PRSV  (40)PSVLLTLFNSGAVEHALNYWIKRDQDVVEVIVLVEQLCRKVTLARTILE
     PPV  (43)PGVLMALFNSGSLEKATQYWITRSHSLAAITSMLSALAAKVSLASTLNA
Consensus     PGVLIALFNSGSLEKA NYWIKRD DLAAV SLLS LA KVSLAKTL
               101                                                150
    LuMV(101)  LDLMQRQAEVLHNLISQSNNHDVSRVVILELLQVMMERNETDSTLLSAG
    PRSV  (90) FNEIRQNARDLHELMDRNNKPWISYDRSLELLSVYANSQLTDEGLLKQG
     PPV  (93) MSVIDEHAAVLYDSVFVGTQPYASYMMAVKTLERMKARTESDHTLNDLG
Consensus      L LI   NA VLHDLI   NNPWISYMMALELL VM  RNETD TLLG
                151                                                200
    LuMV(151)  YMLKNTSSVTLEKSYLKELKEQWDALGYVGKLRHACYLLRHTRKSQPEL
    PRSV(140)  STLDPRLREAVEKTYATLLQEEWRALSLFQKLHLRYFAFKSQPSFSEYL
     PPV(143   SVLRQATPHLVEKSYLQELEQAWKELSWSEKFSAILESQRWRKHIPKPF
Consensus      FSMLKN S   VEKSYL EL E WKALSW  KL    FA R  K L
               201                                                250
    LuMV(201)  PQGEIDLKGIYSVSPRSCLAMTKAHIFNQAHKMKQCAKSLINRARSGTL
    PRSV(190)  PKGRADLKIVYDFSPKYCVHEVGKAFLLPVKAGAKIASRIINGCGAFIR
     PPV(193)  PKDGADLGGRYDISVRSLLGNQYKRLRDVVRWKRDDVVCYTYQSMGKLF
Consensus      PKG ADLKGIYDISPRSCLA   K I   VK  K  A   IIN A A I
               251                                                300
    LuMV(251)  AMCSCINSLVPDFVKLANVLLVMCLMLQLYNTVRMMIKEMHHMKQAKAI
    PRSV(240)  SAAKGCAYIFKDLFQFVHVVLVLSILLQIFRSAQGIATEHLQLKQAKAE
     PPV(243)  KAIGISPSFLPSTLKMLDMLIVFSLLLSIGATCNSMVNEHKHLKQLAAD
Consensus      KAA     SILPD LKLL VLLVLSLLLQIF T N MI EHHLKQAKAD
                301                                                350
    LuMV(301)  ESNERFSRLKKIYGRELMESDPPKPPKFNDFLEEVKRIAPELYDDALL
    PRSV(290)  VERQKDFDRLEALYAELCVKSGEQPTTEEFLDFVMEREPRLKDQAYNL
     PPV(293)  REDKKRFKRLQVLYTRLSEKVGCTPTADEFLEYVGDENPDLLKHAEDL
Consensus      LE NKRF RL  LYARL MKSG    PT DEFLEFV D  PDL D AL
```

Fig.9 Comparison of conserved motives of LuMV P3 to those of other close related potyviruses. The conservative leucine marked in color. The following viruses were included in the analysis: papaya ringspot virus (PRSV), plum pox virus (PPV)

2.2.6 P1 protein

P1 protein was called as mysterious protein because its functions are not clear and sequence encoding the P1 protein, showing a high variability. It was proposed for this protein a role in virus movement and replication. It was found along actin filaments with another membrane-associated potyviral protein 6K which containing replication vesicles (Cui et al., 2010).

Despite of high variability, conserved motives were identified upstream of the quite well conserved protease domain located near P1 C- terminus. The LuMV P1 conservative motives have been illustrated by (Fig. 10) with the % of highest amino acids identities was no more than 26%. The resulting protein is 419 aa long and shared about 26% aa sequence identity with that of *Pea seed-borne mosaic virus,* PSbMV. The serine protease catalytic residues H_{370}-D_{379}-S_{413}-G_{438} are perfectly conserved in the C-terminal part of the P1 protein. The cleavage site between the P1 and the HC-Pro protein was identified 24aa downstream from the conserved G_{438} in the MIHY/SD motive. The resulting protein is 419aa long and shared about 26% aa sequence identity with that of *Pea seed-borne mosaic virus*, PSbMV as revealed by BLAST search. P1 contain a motive characteristic of a serine proteinase H-G-X-D-X-G-X-S-G which have been presented in our LuMV P1 as well by comparison to close potyviruses, this motive is located 48 residues from the cleavage site and 19 residues from the conserved L (F)-I-V-R-G (Sarkisova et al., 2011). It was shown and discussed previously by authors (Aleman et al, 1996; Johansen et al., 1991; Verchot et al., 1992), these both motives involved in the autoproteolytic activity.

```
               601                                          659

     LuMV (359) ENEYVKINTIHECGPRKQIDVCTNRGIDALRIIWSNKIQKKRVIPTCEM
      PVY (181) GAYSAAVRTAHMMGLRRRVDFRCDMWTVGLLQRLARTDKWSNQVRTINI
     JYMV (224) KKVYLKINTKHEEGFNKARDVVMDNFTQRLLELMITRTSGNNRHSVQNI
     KoMV (227) RSRYLRFATKHHEGRRSQRDMPVDHSTSSIQNEAVAVSAFKQPLARGI
      LMV (334) GVRLVKLKTAHEEGHRRKVDIRIPNGLRSIVTRISARGGWHKTWTDSEL
     NYSV (214) SRTYAQVDVHHLSGKLRKTDMDNNTFVNDCLDVMAKITAGNQPHTTEEI
    PanVY (201) EGLKVITKKYHELGVRARKDVLCDEFQQECIRTISSSILKKRVFHENVA
      PPV (205) RLTYFRTHVRHLDGSKPRYDLVLDEATKKILQLFANTSGFHHVHKKGEV
    ScaMV (108) GTIYARAQVKHLQGRRQRRDFESNPALDIWVDILMRRTVGRRTHKTNSI
    SPFMV (561) GANYLFLHLKHMEGLRESVDLRIHTTTQNLVLQAAKVGAWKRTVKTTML
     TuMV (259) RTKFARIQVAHLLGKRAQRDLLTGMEENHFIDILSKYSGNKTTINPGVV
      YMV (196) HLGSCYVRTKHMDGARKKIDLTLNERQEEMLFSLAKG-SRAPIPLSHL
Consensus         YLKV TKHLEG RKR DL I      IL ILAK  SA K  I   I
               651                                           700
     LuMV (409) PGFSGVVVPVANVKGWVGRTKNDILIVRGSNYNQIIDSRVRVTKFVKDK
      PVY (231) RGDSGVILNTKSLKGHFGRSSGGLFIVRGSHEGKLYDARSRVTQSILNS
     JYMV (274) PGHSGFVLNRETLCGTQSRAYGRVFIVRGNHEGKLYDARIKLSQTIRRK
     KoMV (276) TGDSGLAIDANTLGKVGRTFSYGYTIIRGECEGKIFDARSKVTKSIALR
      LMV (384) PGSSGYVLNSSKIIGEFGLRRHSIFVVRGRVYGKIIDSQSKVTHTLTHR
     NYSV (264) FGWSGMVLRTDKLLGKRTKSSLKSFVVRGKDGSRLVDARTRTSYLRRMR
    PanVY (251) KGYSGIVLPQQSVQGKIGRCINGLFIIRGRHEGILLNSLSKLTESVTLR
      PPV (255) PGMSGFVVNPMNLSDPMQVYDTDLFIVRGKHNSILVDSRCKVSKKQSNE
    ScaMV (158) AGWSGFLLNASKLIGRQSTHRGNTFVVRGKCADTLFDARVRMTYDAMLN
    SPFMV (611) KGSSGLVLNPDKLLGPRGHAPHGMLVVRGALRGVLYDARMKLGRSVLPY
     TuMV (309) AGWSGIVVGNGILTQKRSRSPSEAFVIRGEHEGKLYDARIKVTRTMSHK
      YMV (245) TGDSGRIILNP-CDTAYSRYKREGLIVRGEHEGKIYDARVKVTRSVAST
Consensus      T G SGIVLN   L G  GR    LFIVRG HEGKLYDARVKVTKSV R
```

Fig.10 Features conserved in P1 protein of the genus *Potyvirus* (partial aa sequence of P1- C terminal part). The following viruses were included in the analysis: potato virus Y (PVY), Japanese yam mosaic virus (JYMV), Konjac mosaic virus (KoMV), lettuce mosaic virus (LMV), narcissus yellow stripe virus (NYSV), panax virus Y (PanVY), plum pox virus (PPV), scallion mosaic virus (ScaMV), Sweet potato feathery mottle virus (SPFMV), turnip mosaic virus (TuMV), yam mosaic virus (YMV)

2.2.7 Membrane associated protein (6K), viral genome-linked protein (VPg) and cylindrical inclusion (CI) protein

The 6K protein is known as membrane associated protein that can form vesicles and it is connected with the endoplasmatic reticulum where it participate in virus replication or RNA translation as was discussed previously (Cotton et al., 2009; Wei et al., 2010a). The each of two resulting LuMV 6K proteins was 53 aa in length with (Mr.) 6 kDa (Sarkisova et al., 2011). The aa % identities for LuMV 6K was in the range 57 to 67%. The *Papaya leaf distortion mosaic virus*, PLDMV showed the highest % of homology and other *Lettuce mosaic virus*, LMV, *Japanese yam mosaic virus,* JYMV and *Plum pox virus,* PPV shared the high similarity *to Lupine mosaic virus*, LuMV.

The VPg is known by bounding to the 5'-termini of genomic RNAs and interacting with the translation initiation (Puustinen et al., 2004; Roudet-Tavert et al., 2007). To compare LuMV VPg the multiply alignment was done and phylogenetic relatedness (data not shown) estimated which revealed that aa % of entire amino acids identities lies in range between 53 to59% and 67-74% for only conserved amino acids (Tab. 7). It was done by using the DNASTAR software package (v.8.0.2), (Lasergene, Madison, WI, USA). *Japanese yam mosaic virus,* JYMV, *Narcissus yellow stripe virus,* NYSV, *Plum pox virus*, PPV, *Turnip mosaic virus*, TuMV and S*callion mosaic virus,* ScaMV were shown to be closely related to our *Lupine mosaic potyvirus*, LuMV (Tab. 7, 8.). The highest % of homology Lupine *mosaic virus*, LuMV on nt level was 61% to *Plum pox virus*, PPV (Tab.8). The LuMV VPg protein is resulted in 193 aa and with (Mr.) 21.4 kDa (Sarkisova et al., 2011).

Tab. 7A sequence identity comparison (%) of entire amino acids (above the diagonal) and conserved only (bellow the diagonal) of VPg between LuMV and other close related viruses

	LuMV	JYMV	NYSV	PPV	TuMV	ScaMV
LuMV		54	56	59	57	53
JYMV	67		67	58	65	67
NYSV	68	80		64	68	71
PPV	74	76	81		62	60
TuMV	68	81	80	76		67
ScaMV	70	82	84	78	81	

***full viral names including abbreviations in Tab.1**

Tab. 8 Nucleotide sequence identity comparison of VPg regions between LuMV and other close related viruses

	LuMV	JYMV	NYSV	PPV	TuMV	ScaMV
LuMV		55	59	61	57	57
JYMV			63	56	64	63
NYSV				60	66	68
PPV					61	60
TuMV						67
ScaMV						

***full viral names including abbreviations in Tab.1**

DNA and RNA helicases are enzymes are required for the process of unwinding of nucleic acid duplexes (Gorbalenya et al., 1993). CI protein has been shown since long time, it also functions as RNA helicase and involved in virus replication (Lain et al., 1990). The few conservative domens were determined in LuMV Cl that is in correspondence to other potyviruses (Fig.11). BLAST search revealed that *Konjac mosaic virus*, KoMV, *Panax virus Y*, PanVY, *Plum pox virus*, PPV, *Japanese yam mosaic virus*, JYMV, *Narcissus yellow stripe vir*us, NYSV, *Scallion mosaic virus,* ScaMV, and *Turnip mosaic virus*, TuMV. The % of homology on aa level lies in range between 58-62%. The highest entire amino acids identities LuMV Cl motive shared to *Konjac mosaic virus,* KoMV and *Plum pox virus*, PPV with 61 and 62% and 78% for only conserved amino acids, respectively (Tab. 9.). The highest nt level of identities was 64% to *Konjac mosaic virus,* KoMV (Tab.10). The conservative motives have been recognized in Cl region as illustrated by (Fig.11).The LuMV Cl protein had weight of (Mr.) 71.5 kDa with 644 aa in length (Sarkisova et al., 2011).

```
               1                                                  51
   LuMV (1)    LDDIESLDDDKKLTIDFEIDGKQAHGCTYKEMSFSKWWGIQMDNNRAVP
   KoMV (2)    LDEEDTIQFEKWQTIDFELAGNELMAPSIQEMSFQQWWEHQLQNNRTIP
   PPV  (1)    LDDIEDILDEKKLTVDFVLQSNEVAPTVPFDSTFEKWWMNQLETGNVIP
Consensus      SLDDIESIDEKKLTIDFEL GNEL A S EMSF KWW QLDNNRIP
               52                                                101
   LuMV (51)   YRTEGFFMEFTRSTATQVANEIAMGAHKDILLRGAVGSGKSTGLPSQLA
   KoMV (52)   YRTEGHFMEFTRARASIVANEIAHSDHKDILLRGAVGSGKSTGLPYHLS
   PPV  (51)   YRTEGHFLEFTRENAAHIANEVMHGSHQDILIRGAVGSGKSTGLPFHLS
Consensus      HYRTEGHFMEFTRA AS ANEIAHGAHKDILLRGAVGSGKSTGLPFHLS
               102                                               151
   LuMV(101)   KGRVLMIEPTRPLAENVCKQLRGDPFHLNATLRMRGMSTFGSSPITIMT
   KoMV(102)   KGRVLLLNPTRPLAENVWRQLKADPFYMNPTLRMRGTSMFGSSPVHIMT
   PPV (101)   KGHVLLIEPTRPLAENVCKQLRGQPFNVNPTLRMRGMSTFGSTPITVMT
Consensus       GRVLLIEPTRPLAENVCKQLRGDPFHLNPTLRMRGMSTFGSSPITIMT
               152                                               201
   LuMV(151)   GYALHYLAHNRQTISDFDFIIFDECHVNDSSAMAFRCLLYDQQYEGKII
   KoMV(152)   GYALHYLANNQNIICEYDFILFDECHVLDASAMAFRSLLAEYDFRGKII
   PPV (151)   GYALHFLANNPTYLDNYKCIIFDECHVHDASAMAFRCLLSEYSYPGKIL
Consensus      GYALHYLANN N I DYDFIIFDECHV DASAMAFRCLLAEY Y  KII
               202                                               251
   LuMV(201)   VSATPPGREVEFTTQYPVKILIEESLTFAQFVLNQDTKVNSDIVQYGNN
   KoMV(202)   VSATPPGREVEFTTQFPVDIKIEENLTFDQFVQGQGTKANCDVLQHGNN
   PPV (201)   VSATPPGHEVEFKTQKEVKVIVEESLSFQQFVSNLGTGCNSDILKHGVN
Consensus()    KVSATPPGREVEFTTQFPVKIIIEESLTF QFV NQGTK SDILQHGNN
               252                                               301
   LuMV(251)   LVYVSSYSEVDNLSKLLIEKKYKVTKVDGRTMKLGNVNIETTGTDSQKH
   KoMV(252)   LVYVASYSEVDMLSKKLSEKNYSVTKVDGRTMKVGSVNIQTHGTTEKPH
   PPV (251)   LVYVASYNEVDTLSKLLTDRSFKVSKVDGRTMKVGNVEIPTSGTQAKPH
Consensus      ILVYVASYSEVD LSKLLSEK YKVTKVDGRTMKVGNVNI TSGTAKPH
```

```
            302                                             351
   LuMV(301) IVATNIIENGVTLDVDVVVDFGLKVSPQLDPDNRMISYRKEAISYGERI
   KoMV(302) VVATNIIENGVTLDVEVVVDFGLKVVAQLDCDNRCMVYTKKSVSFGERI
   PPV  (301) VVATNIIENGVTLDIDVVVDFGLKVVPVLDIDNRLVRYTKKSISYGERI
Consensus    FVVATNIIENGVTLDVDVVVDFGLKVVPQLD DNRLIYTKKSISYGERI
            352                   380
   LuMV(351) QRLGRVGRHKSGVALRIGHTERGLVQIPS
   KoMV(352) QRLGRVGRQKPGMALRIGTTEKGLMEIPS
   PPV  (351) QRLGRVGRNKPGAALRIGFTEKGLTQIPP
Consens(352)QRLGRVGRNKPGMALRIG TEKGLMQIPS
```

Fig.11 Conserved motives LuMV Cl (partial sequence) to that of other close related potyviruses. Conservative domains were detected the alignment partial sequences of the CL motive of closest related viruses. The following viruses were included in the comparison: konjac mosaic virus (KoMV), plum pox virus (PPV)

Tab. 9 Sequence identity comparison (%) of entire amino acids (above the diagonal) and conserved only (bellow the diagonal) of Cl regions between LuMV and other close related viruses

	LuMV	PPV	JYMV	NYSV	ScaMV	TuMV	KoMV	PanVY
LuMV		**62**	**59**	**59**	**58**	**58**	**61**	**59**
PPV	**78**		**62**	**61**	**61**	**62**	**62**	**63**
JYMV	**76**	**78**		**69**	**70**	**69**	**60**	**59**
NYSV	**75**	**78**	**84**		**73**	**71**	**59**	**58**
ScaMV	**76**	**78**	**84**	**86**		**70**	**60**	**60**
TuMV	**75**	**77**	**84**	**86**	**84**		**58**	**58**
KoMV	**78**	**79**	**76**	**74**	**75**	**75**		**60**
PanVY	**76**	**78**	**76**	**76**	**76**	**76**	**77**	

***full viral names including abbreviations in Tab.1**

Tab. 10 Nucleotide sequence identity comparison (%) of Cl regions between LuMV and other close related viruses

	LuMV	PPV	JYMV	NYSV	ScaMV	TuMV	KoMV	PanVY
LuMV		**61**	**61**	**60**	**59**	**59**	**64**	**61**
PPV			**60**	**62**	**61**	**60**	**61**	**62**
JYMV				**66**	**66**	**66**	**60**	**60**
NYSV					**66**	**67**	**60**	**58**
ScaMV						**64**	**58**	**60**
TuMV							**58**	**59**
KoMV								**60**
PanVY								

***full viral names including abbreviations in Tab.1**

2.2.8 3 - 5' UTRs

The 5'- and 3' of LuMV was comprised of 130 and 212 nt, respectively and both terminal regions were amplified by using PCR and nested PCR (5-RACE system protocol), then products were cloned with pJET or TOPO vectors and sequenced (Sarkisova et al., 2011). Primers used in amplification were from 5-RACe kit, and oligo d(T) primer and specific ones which have been designed from nucleotide sequences flanked untranslated regions. They were used in combinations: d(T) with Spoty- potyvirus specific primer for CP of potyviruses, AAP and UAAP primers with CIRev primer for 5'RACE(System for Rapid Amplification of cDNA Ends (Invitrogen, USA) with CI Rev primer which was used from article Ha et al., (2008). Both ends were examined for their potential secondary structures using the MFOLD web service (http://mfold.rna.albany.edu/?q=mfold). The detailed analysis of untranslated regions is bellow.

The 5'UTR leader sequence of LuMV ssRNA consisted of 130 nt in length as mentioned above that was in correspondence with other the potyviruses. Actually, this genome fragment of potyviruses is varied in the length range from 8nt to 205nt. The 5' UTR was rich in adenine and uracyl content - 58 and 28%, respectively but displays a very low percentage of guanine residues 10% and cytosine 34%, respectively (Sarkisova et al., 2011). This feature as was shown to be common not only by other potyviruses and by many plant virus 5' leader sequences (Galie et al.,

1987). 5'-UTR was examined for the presence of secondary structures; however, folding of the 5'UTR using the mfold web server did not predict stable secondary structure. Unlike of other potyviruses, for example, it was shown on Turnip *mosaic virus*, TuMV RNA (Nicolas et al., 1992). Overall homology of the LuMV 5'UTR with those of other potyviruses based on alignment of 5'UTR regions ranges on nt level from 58 -59% to 69% identities with *Lily yellow stripe virus*, LYSV, *Clover yellow vein virus*, CLYVV and *Potato virus A,* PVA, respectively. Thus the 5'UTR was poorly conserved except for two highly conserved regions of unknown function, box “a” and box “b” which were observed as was discussed previously by Lain et al., (1989) and present in other potyviruses. In the LuMV 5'UTR, box “a” (ATACAACTT) comprising nt 21 to 31nt slightly different from the consensus sequence (ACACAACAU) while box “b” (UCAAGCA) includes nt 65 to nt 71 and is identical to the consensus sequence (Sarkisova et al., 2011; Aleman et al.,1996).

LuMV 3'UTR comprised 212nt in length and it is adenine and uracyl content was 49 and 88% that was higher to that of guanine and cytosine – 40 and 35%, respectively (Sarkisova et al., 2011). In (Fig.12) showed the predicted secondary structure of the 3’UTR: major stem-loop structures with free energy dG= -37kcal/mol (Fig. 12). Similar step-loops structures have been identified in different potyviruses and similar secondary structures are thought to be responsible for the replication, regulation of transcription and symptom expression (Matthewes, 1991; Rodriduez-Cerezo, 1991). Comparison of LuMV 3'UTR nucleotide sequence with other fully sequenced potyviruses displays to *Basella rugose mosaic virus,* BaRMV, *Daphne mosaic virus*, DapMV, *Pea seed-borne mosaic virus*, PSbMV and *Panax virus Y*, PanVY the highest % of identities 30-63 % (Sarkisova et al., 2011) what is slight different from previous data (39-53%) by Frenkel et al., (1989).

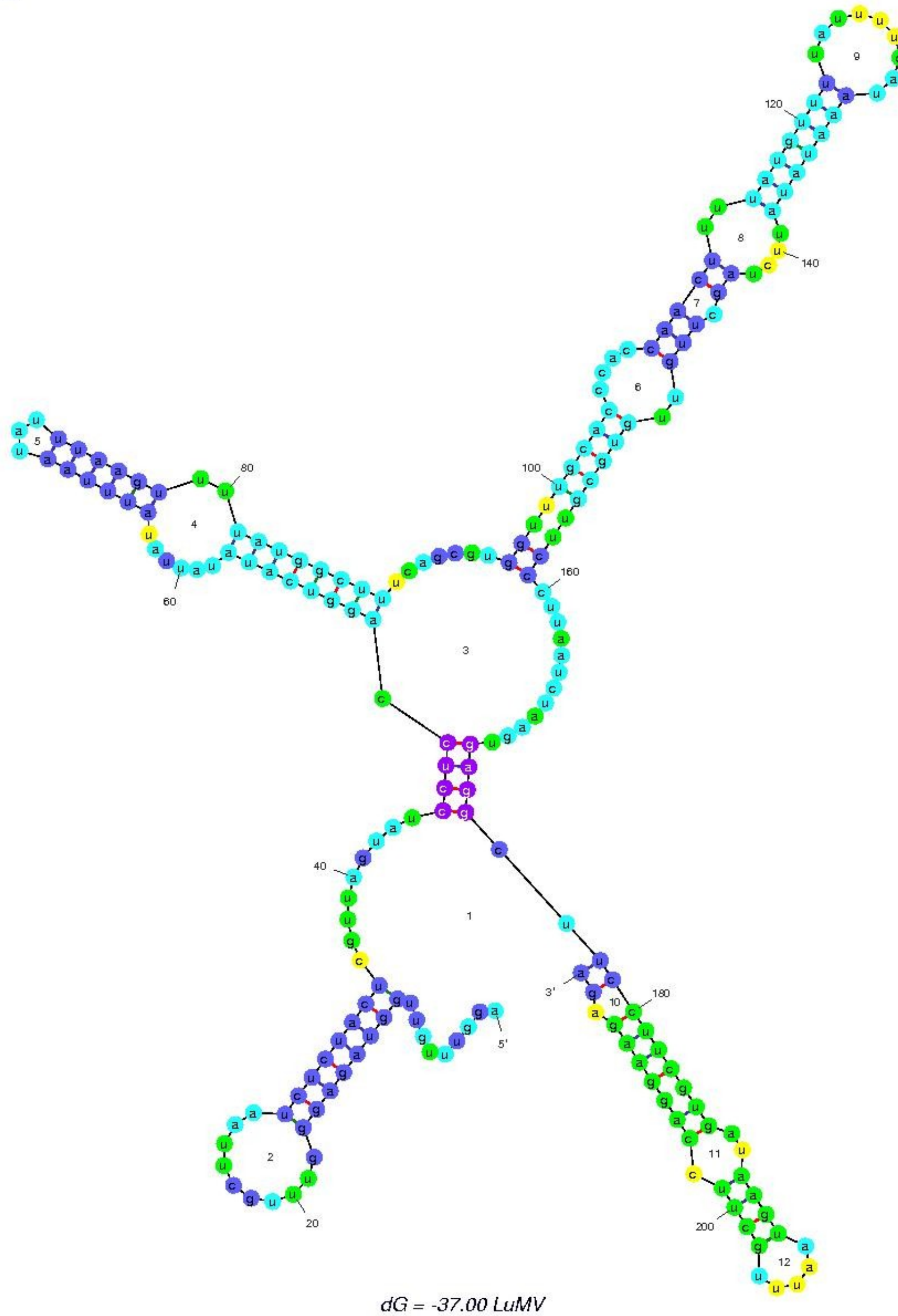


Fig.12 The potential secondary structure of the 3'- end of (+) strand LuMV from *Lupinus polyphyllus.* The putative folded structure of the 3'- end was obtained from the MFOLD program (Zuker et al., 1999)

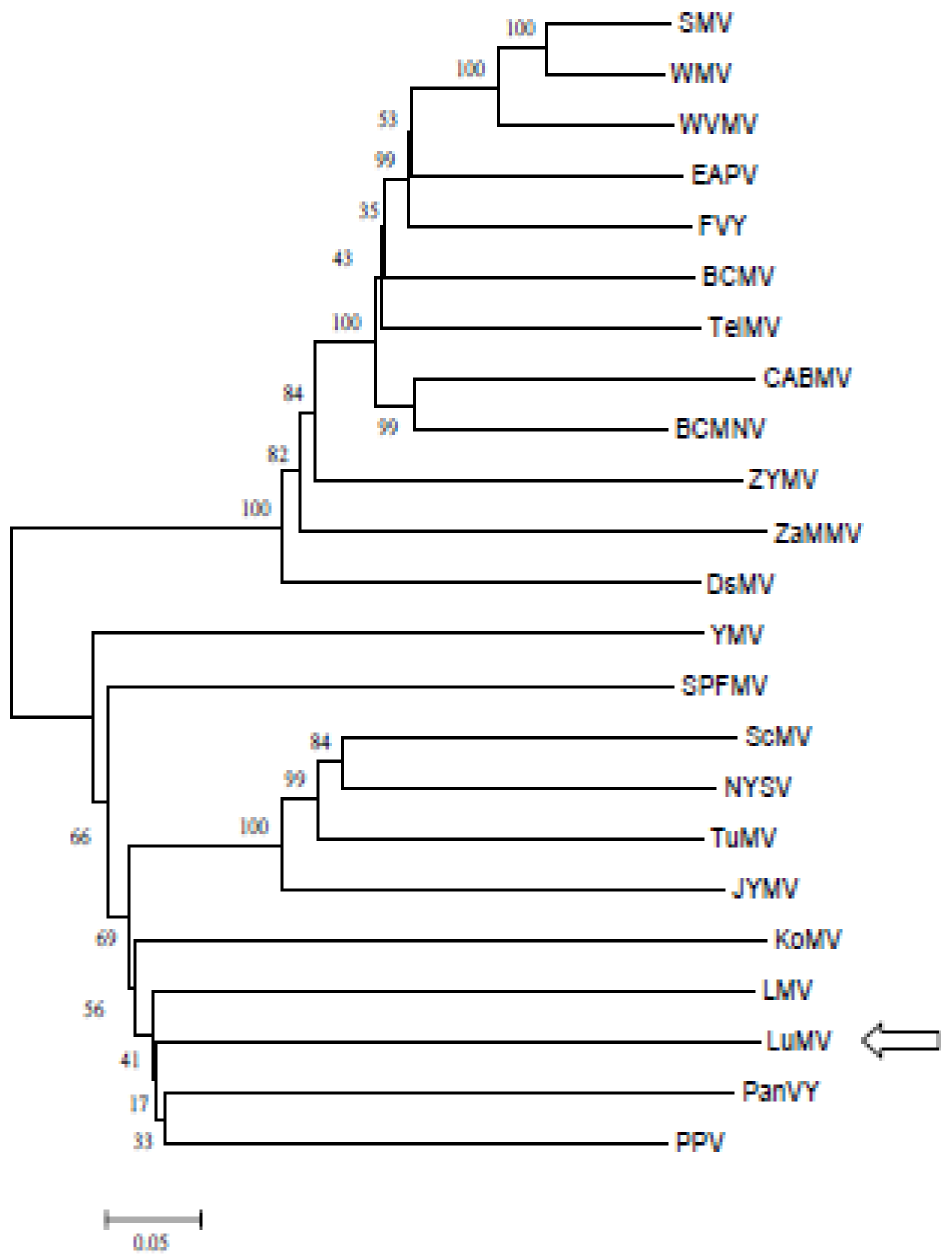

SMV
WMV
WVMV
EAPV
FVY
BCMV
TelMV
CABMV
BCMNV
ZYMV
ZaMMV
DsMV
YMV
SPFMV
ScMV
NYSV
TuMV
JYMV
KoMV
LMV
LuMV
PanVY
PPV
100
100
53
99
35
43
100
84
99
82
100
84
99
100
66
69
56
41
17
33
0.05

Fig.13 Bootstrapped neighbor-joining tree obtained using the complete polyprotein sequences of LuMV and others 22 closest potyviruses. Accession numbers are indicated in brackets in the legend. The scale bar represents a genetic distance of 0.05. The following viruses were included in the analysis: bean common mosaic virus (BCMV; NC_003397), bean common mosaic necrosis virus (BCMNV; NC_004047), cowpea aphid-borne mosaic virus (CABMV; NC_004013), dasheen mosaic virus (DsMV; NC_003537), East Asian passiflora virus (EAPV; NC_007728), fritillary virus Y (FVY; NC_010954), Japanese yam mosaic virus (JYMV; NC_000947), konjac mosaic virus (KoMV; NC_007913), lettuce mosaic virus (LMV; NC_003605), narcissus yellow stripe virus (NYSV; NC_011541), panax virus Y (PanVY; NC_014252), plum pox virus (PPV; NC_001445), scallion mosaic virus (ScMV; C_003399), sugarcane mosaic virus (SCMV; NC_003398), soybean mosaic virus (SMV; NV_002634), sweet potato feathery mottle virus (SPFMV; NC_001841), telosma mosaic virus (TelMV; NC_009742), turnip mosaic virus (TuMV; NC_002509), watermelon mosaic virus (WMV; NC_006262), wisteria vein mosaic virus (WVMV; NC_007216), yam mosaic virus (YMV; NC_004752), zantedeschia mosaic virus (ZaMMV; NC_011560), zucchini yellow mosaic virus (ZYMV;NC_003224)

Tab. 11 Species of genus *Potyvirus* used in this study

Species name		Species name		Species name	
AWMV –	***Algerian watermelon mosaic virus***	**LMoV -**	***Lily mottle virus***	**SrMV –**	***Sorghum mosaic virus***
BaRMV -	***Bassela rugose mosaic virus***	**LYSV –**	***Leak yellow stripe virus***	**SMV –**	***Soybean mosaic virus***
BCMV -	***Bean common mosaic virus***	**MDMV –**	***Maize dwarf virus***	**ScMV –**	***Sugarcane mosaic virus***
BCMNV –	***Bean common mosaic necrosis virus***	**MWMV –**	***Moroccan watermelon mosaic virus***	**SYSV –**	***Shallot yellow stripe virus***
BBrMV -	***Banana bract mosaic virus***	**NDV -**	***Narcissus degeneration virus***	**SuCMoV**	***Sunflower chlorotic mottle* virus**
BYMV -	***Bean yellow mosaic virus***	**NYSV -**	***Narcissus yellow stripe virus***	**SPFMV -**	***Sweet potato feathery mottle virus***
BtMV -	***Beet mosaic virus***	**OYDV -**	***Onion yellow dwarf virus***	**TEV –**	***Tobacco etch virus***
CABMV	***Cowpea aphid-born mosaic virus***	**PanVY –**	***Panax virus***	**TFMV –**	***Thunberg fritillary mosaic* virus**
CaYSV –	***Canna yellow streak mosaic virus***	**PLDMV -**	***Papaya Lea distortion virus***	**TVBMV -**	***Tobacco vein banding mosaic virus***
ChiVMV –	***Chilli veinal mottle virus***	**PPV –**	***Plum Pox virus***	**TelMV –**	***Telosma mosaic virus***
CIYVV -	***Clover yellow vein virus***	**PepMoV –**	***Pepper mottle potyvirus***	**TVMV –**	***Tobacco vein mottling virus***
CSV -	**Cocksfoot streak virus**	**PepSMV -**	***Pepper severe mosaic* virus**	**TuMV –**	***Turnip mosaic virus***
DapMV –	***Daphne mosaic virus***	**PTV -**	***Peru tomato virus***	**VVY –**	***Verbena virus Y***
DsMV -	**Dasheen mosaic virus**	**PSBMV –**	***Pea seed-borne mosaic virus***	**YMV –**	***Yam mosaic virus***
EAPV –	***East Asian Passiflora virus***	**PVV –**	***Potato virus V***	**WMV –**	***Watermelon mosaic virus***
FVY -	***Fritillary* virus *Y***	**PRV –**	***Papaya ringspot virus***	**WPMV-**	***Wild potato mosaic virus***
FreMV –	***Freesia mosaic virus***	**PeMoV -**	***Peanut mottle virus***	**WTMV –**	***Wild tomato mosaic virus***
JGMV –	***Johnson grass mosaic virus***	**PVA –**	***Potato virus A***	**WVMV –**	***Wisteria vein mosaic virus***
JYMV –	***Japanese yam mosaic virus***	**PVY -**	**Potato *virus Y***	**ZaMMV-**	***Zantedeschia mild mosaic* virus**
KoMV –	***Konjak mosaic virus***	**PVMV –**	**Pepper *veinal mottle virus***	**ZYMV –**	***Zucchini yellow mosaic virus***
LMV -	***Lettuce mosaic virus***	**SCMV –**	**Sugarcane mosaic virus**	**LuMV –**	***Lupine mosaic virus***

The all data are presented in this chapter are corresponding to the following papers and short conversation in book of abstracts:

Sarkisova T., Petrzik K. 2009. Analysis of capsid protein sequence revealed new potyvirus found in *Lupinus polyphyllus* in the Czech Republic. Book of abstracts: XVIII Czech and Slovak plant protection conference. 35.

Sarkisova T., Petrzik K. 2009. A new potyvirus identified in Czech Republic. Acta Virol. 53: 143. **(IF: 0,682)**

Sarkisova T., Petrzik K. 2011. Determination of the complete nucleotide sequence of a lupine potyvirus isolate from Czech Republic reveals that it belongs to a new member of the genus *Potyvirus*. Arch. Virol. 156: 167-169. **(IF: 2,111)**

Chapter 3
Results

3.1 Summary of screening phytoplasma- infected for the plasmid presence

To perform this part of the work 37 phytoplasma- infected samples were obtained and subjected to the detection of plasmid presence. In our work, pairs of specific universal primers of our design have been used for the plasmid screening (Petrzik et al., 2011). It aimed beside the detection plasmids to see if there is possibility to use plasmids' presence as a marker for the considering of the phytoplasma infection because plasmid amplification is less time consuming, it can be easier selecting of suspected samples and then, to continue with only positive ones. Up to now, there is limited information of phytoplasmas plasmids except their usual presence in high copy numbers up to 20 per cell and more probably related to increasing adaptation of phytoplasmas in changing environments (Eberhard, 199) less known of their distribution between phytoplasmas groups and subgroups. As was shown recently the plasmids could be associated with transmissibility that was shown for the plasmid DNA from *Clover phyllody* phytoplasma and its rearrangement over relatively short period what was observed in clover plants which were regenerated by tissue culture and they lack of *ssb* protein what caused inability to the insect transmission (Denes et al., 1992). Similarly, results were obtained by Oshima et al., (2001) who determined that plasmids in a non-insect-transmissible line of *Onion yellows* phytoplasma after maintenance by grafting for 2 years, lack this protein as well. Their significant role in adaption to changes of environmental conditions could be explained by evolving more rapidly than genes on chromosomes (Eberhard, 1990).

The presence in few copies per cell – the high concentration could be useful in the detection of phytoplasmas by doing it is more easily. By the way, there are a lot of black spots related to them as, for example, recently found helicase domains in *rep* gene which has shown high homology to eukaryotic helicase so there is a question how it is was acquired? If it is somehow related to the plasmid origin or is a result of their recombination. As well is not much information of their distribution between phytoplasma strains and thus, it is also interested if there is any connection of the plasmid to the certain strain or not. Using molecular tools, phytoplasma infected samples have been examined for the plasmid presence and as

the result – their content was confirmed in the half of testing phytoplasma infected samples.

3.1.2 Detection of plasmids by using specific primers for *rep* genes

During the initial PCR within *rep* gene universal primers 423K8/K9 (Petrzik et al., 2011) generated partial fragments of phytoplasma plasmids. DNA used in amplification was coincidentally isolated from samples of phytoplasma infected frozen material (Tab.12). Fragment's predicted size was around 400 bp in length, which spanned the N-terminal part of the gene; our amplified products a little varied in length (data not shown here). Sequence BLAST search of partial fragment sequences was done and alignments were created by DNASTAR software package (v.8.0.2) (Lasergene, Madison, WI, USA) and used in analysis. Results pointed that our replicase gene partial sequences showed high similarity to pOYW.

Thus, from 37 phytoplasma - infected samples (Tab.12) subjected for the screening in 19 plasmids presence have been found. The multiply alignment of 18 (pPEY from *Pichris echioides yellows* phytoplasma was not used because of lack data from the central part of *rep*) fragments of *rep* gene were used in phylogenetic analysis and showed several clades (Fig.15). Plasmids from different 16S ribosomal groups were not always grouped in distinct clusters, which pointed that not all plasmids have close relationships with their phytoplasma genomes.

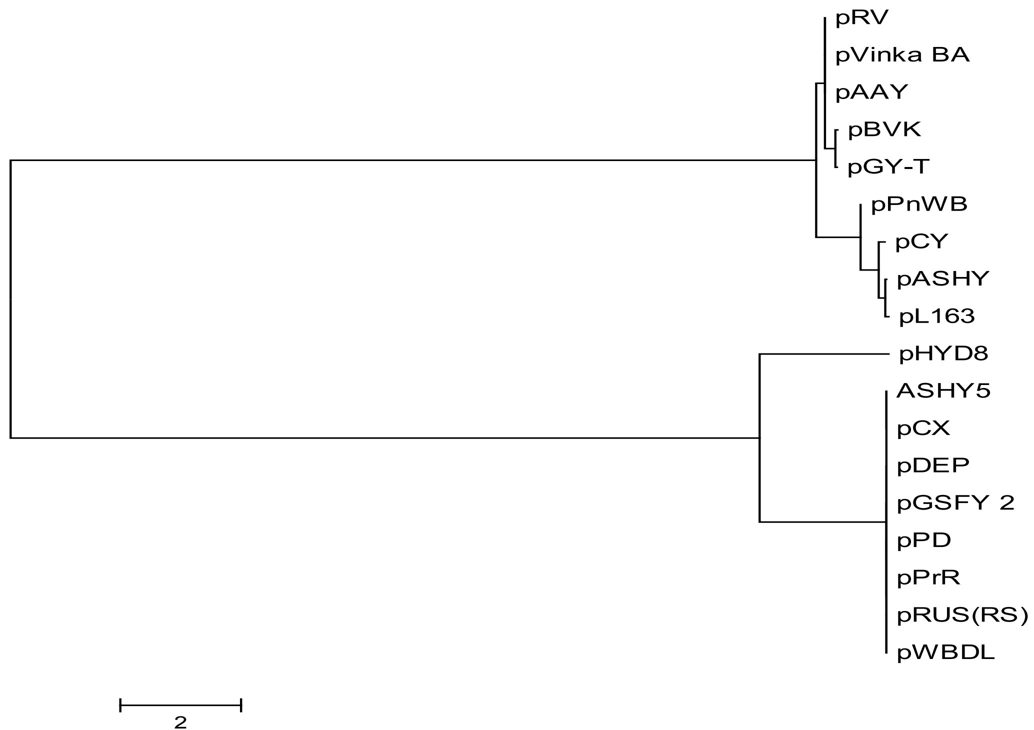


Fig.14 Phylogenetic tree for the plasmid's N- terminal part of *rep* protein gene computed with the maximum parsimony algorithm in MEGA (Tamura et al., 2007) of 18 detected phytoplasmas plasmids; abbreviation in (Tab.12)

3.1.3 Partial sequence analysis of extrachromosomal DNA from *Pichris echioides yellows*, PEY and *Psammotettix cephalotes' flower stunt,* BVK phytoplasma

Two plasmids from plant-pathogenic mollicute "*Candidatus* Phytoplasma *Pichris echioides yellows,* PEY and *Psammotettix cephalotes' flower stunt*, BVK associated with phytoplasma classified to the 16Sr IX and 16SrXI-C subgroup, respectively, were subsequently taken to complete sequencing. These two plasmids were chosen in order to obtain the data of plasmids' organization from the groups where they were not found or described previously. There are several known plasmids belonging to phytoplasma of *Candidatus Aster yellows* group. Total DNA was isolated from frozen samples of infected plants derived from EU-phytoplasma collection (Sarkisova et al., 2012).

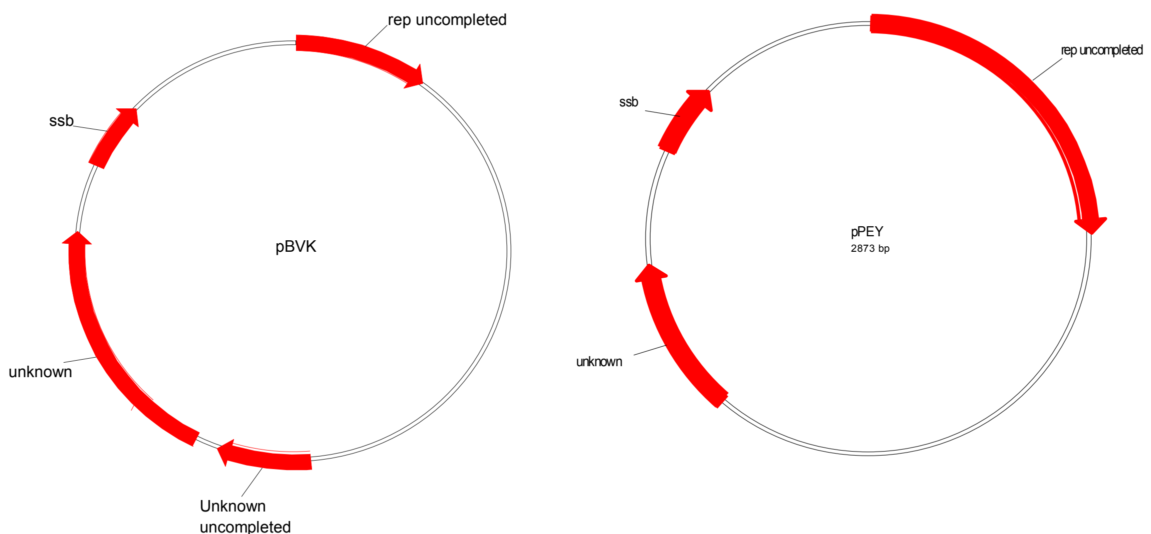
PCR with specific primers 423K8/K9 for the replicase gene (Petrzik et al., 2011) were used in amplification these two plasmids, however, the amplified segment

was obtained only in phytoplasma from the insect host *Psammotettix cephalotes*, BVK. To continue sequencing additional primers' pair for *rep* gene has been designed based on the sequence information and multiply alignment of avaible phytoplasmas' plasmids from GenBank. Thus, several pDNA-specific primers were used to amplify other parts of *rep* gene in several combinations. To extend sequencing in both directions by following four primers' pair: 253/ 426C7, 254/C8 (Tab.1) (Sarkisova et al., 2012). Fragments of *rep* gene were amplified only in pPEY by primer' pairs and consisted from two fragments for pPEY: 250 bp and 586 bp with the gap between them and. There was no possibility to overlap fragments because of luck sequence data in combination 423K8/K9. BLAST search of partial fragment sequences and multiply alignment was done by using the DNASTAR software package (v.8.0.2) (Lasergene, Madison, WI, USA) Molecular analysis within second fragment of rep gene of pPEY did possible to determine II and III conservative motives: R-II ILHDKDIYQN, R-III PHWHIYLRFN and KNIAQW, which presented in plasmids was shown previously. The result was pointed on the structure of *rep* gene of plasmids, which belong to pLS1 family the representatives of which have RC replication strategy as was described by (Solar et al., 1993; Khan, 1997; Oshima et al., 2001). Unfortunately, the lack of sequence data for the whole *rep* gene did not give to recognize all conservatives' motives, which are more probably presented because of similar plasmid organization.

The other producing products varied in size and spanned complete unknown hypothetical protein and *ssb* for both plasmids and including N-terminal part of other unknown hypothetical protein gene for pBVK only (Fig.15) with unknown functions. It consisting from fragments 400 and 1043 bp for pPEY and pBVK, respectively (Sarkisova et al., 2012).To amplify *ssb* gene PCR was done by using primers of our design: 226/503T0 (Tab. 1) which spanned *ssb* gene completely. This gene presence was shown to be necessity and always present in several completely sequenced phytoplasma plasmids. To further amplification of these plasmids regions, few primers were designed based on alignment of available plasmids' sequences in GenBank: 225,307, 315, and 136, which could completely spanned two unknown hypothetical proteins. This piece of plasmids was successfully amplified but obtained sequences were shorted than was predicted. PCR products of *rep*, *ssb*, unknown hypothetical proteins sequences were analyzed

by using the DNASTAR software package (v.8.0.2) (Lasergene, Madison, WI, USA).

Based on obtained data, pBVK and pPEY plasmids showed similarity to pOYW by phylogenetic analysis of N-terminal part of *rep gene*. To see the position of our two plasmids toward of pOYW was proposed because pOYW was first plasmid where found helicase domain with relatedness to gemini- and eukaryotic viruses RCR initiator of Rap which was analyzed by Koonin et al., (1992), Khan, (1997). In addition, phylogenetic analysis of pPEY and pBVK, pOYW to others representative of pLS1 plasmids' family was done to estimate relatedness (concluded after finding of conservative motives in fragments of *rep* gene), it is illustrated by (Fig. 19, 20). Partial sequences of both plasmids have been submitted to GenBank under accessions numbers: KC545788 - pBVK *rep* gene, KC505535- pPEY *rep* gene with conservative motives, KC545789- unknown and ssb proteins of pPEY, KC545790- pBVK unknown, *ssb* and N-terminal part of another unknown protein.



pBVK phytoplasma from *P. cephalotes* **pPEY Pichris echioides yellows**

Fig.15 Schematic representation of both p PEY and pBVK phytoplasmas plasmids (in color sequenced fragments)

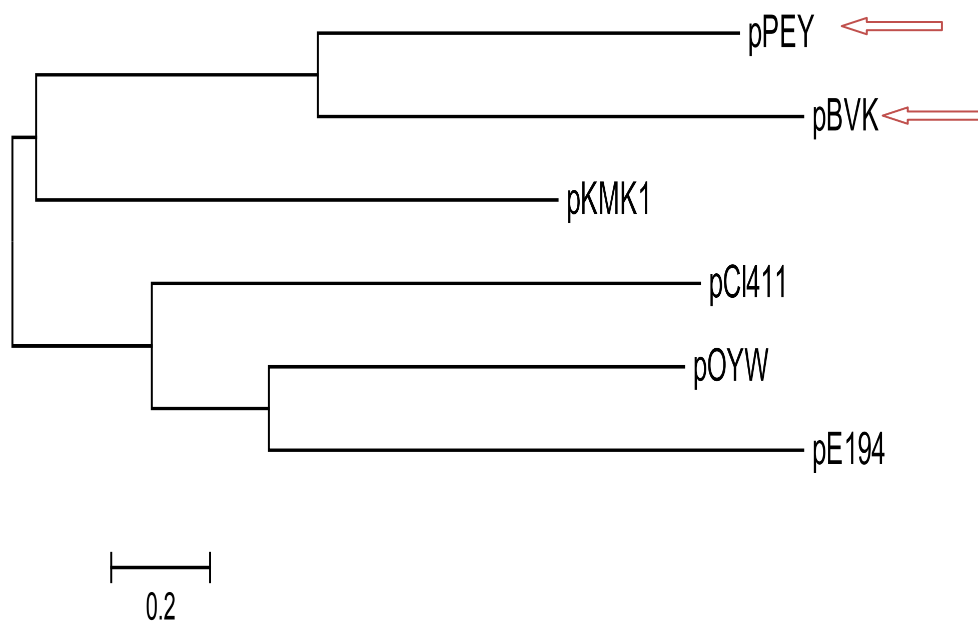


Fig.16 A neighbor-joining (NJ) tree for the amino acid sequences of plasmids family pLS1. Genetic relatedness of both pPEY and pBVK was analyzed by comparison of partial nt sequences of both our phytoplasma plasmids in comparison to others representatives of family pLS1:pKMK1- from *Mycoplasma mycoides*, pCI411- *Leuconostoc lactis*, and pOYW - plasmid from *Onion yellows* phytoplasma, pE194- *Staphylococcus aureus*

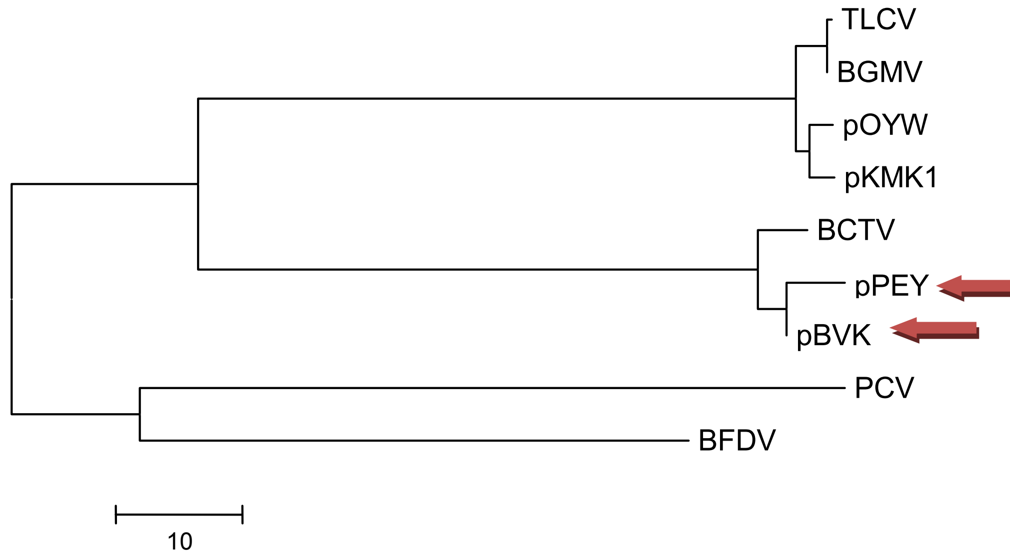


Fig.17 A maximum likelihood tree was drowned for the partial amino acid sequences of the N-terminal region of rep gene of plasmids belonging to pLS1 family including pPEY and pBVK, and eukaryotic viruses. Genetic relatedness was analyzed by comparison conserved motives in the Reps regions of both plasmids and Raps sequences of eukaryotic viruses which contain motives similar to those found in the analogous domains of Rep from the pLS1 and pUB110 plasmid families: geminiviruses (beet curly top virus, BCTV; tomato leaf curl virus, TLCV, bean golden yellow mosaic virus, BGMV), circoviruses (porcine circovirus, PCV, beak and feather disease virus, BFDV), pKMK1- from *Mycoplasma mycoides*, pOYW- *Onion yellows* phytoplasma

Tab.12 A list of phytoplasma strains used for plasmids screening

N	Strain acronym	Strain	Country	groups
1	HYD8	Hydrangea phyllody	*Bordeaux*, France	16SrI
2	CHRY	Chrysanthemum yellows	*C. frutescens*, Berlin, Germany	16SrI-A
3	DAY	Dwarf aster yellows	Clover, CA, USA	16SrI-B
4	GLAWC	Gladiolus witches' broom	Gladiolus with *M. quadripunctulatus*	16SrI-B
5	PrR	Primula red	*Bordeaux*, France	16SrI-B
6	RV	Rape phyllody	Colza, Montfavet, France	16SrI-B
7	L163	Lettuce yellows	Lettuce, Liguria, Italy	16SrI-B
8	CY	Chrysanthemum yellows	*Chrysanthemum frutescence*, Lugaria, Italy	16SrI-B
9	AAY	American aster yellows	Aster, Florida, USA	16SrI-B
10	GY-T	Grapevine yellows	*V. vinifera*, TO, Italy	16SrI-B
11	KAY	Koolsard aster yellows	Bordeaux, France	16SrI-B
12	AKV	Aquilegia virescence	*Aquilegia alpine*, Germany	16SrI-B
13	AY-11	Aster yellows	Maryland , USA	16SrI-B
14	CVT	Catharanthus virescence	Periwinkle, Thailand	16SrI-B
15	PPT	Potato purple top	Potato, France	16SrI-C
16	A-AY (ACLR)	Apricot chlorotic leafroll	Apricot, Valencia, Spain	16SrI-F
17	CVB	Catharanthus virescence	Leafhopper Dossenheim, Germany	16SrI-F
18	PnWB	Peanut witches'	Lee USA	16SrII-A
19	SEPT	Sesame phyllody	Sesame, Thailand	16SrII-A
20	WBDL	Witches broom disease	Lime, Oman	16SrII-C
21	CX	X disease	California, USA	16SrIII-A
22	WX	Western X	Peach, CA, USA	16SrIII-A
23	GVX	Green Valley X	Cherry, California, USA	16SrIII-A
24	JR 1 (PoiBI)	Poinsettia branching factor	Poinsettia, USA	16SrIII-H
25	PEY	Pichris echioides yellows	*Pichris echioides*, PZ, Italy	16SrX
26	EY-C	Elm yellows	*Ulmus americana*, NY, USA	16SrV-A
27	RUS (RS)	Rubus stunt	*Rubus sp.*, Italy	16SrV-E
28	PWB	Potato witches' broom	Potato, USA	16SrVI
29	ASHY5	Ash yellows #5	*Fraxinus americana*, USA	16SrVII-A
30	ASHY	Ash yellows	*Fraxinus Americana*, NY, USA	16SrVII-A
31	PLN V6	Plum leptonecrosis	*Prunus salicina*, Italy	16SrX-B
32	GFSY 2	German stone fruit yellows	Apricot, Germany	16SrX-B
33	PD	Pear decline	Pear 207/86t, Germany	16SrX-C
34	BVK	Flower stunting	Psammotettix cephalotes, Germany	16SrXI-C
35	Vinca BA	Grapevine yellows	*V. vinifera*, BA, Italy	16SrXII
36	GY-US	Grapevine yellows	*V. vinifera*, BO, Italy	16SrXII
37	DEP	Deperissment du lavandin	Lavandin, France	16SrXII-A

The all data are presented in this chapter are corresponding to the short conversation in the book of abstracts:

Sarkisova T., Petrzik K. 2012. Sequence analysis of two plasmids from *Pichris echioides yellows* and BVK phytoplasma from *Psammotettix cephalotes*. Book of abstracts: XIX Czech and Slovak plant protection conference. 154-155. ISBN 978-80-552-0838-1

Petrzik K., Sarkisova T., Curnova L. 2011. Universal primers for plasmid detection and method for their relative quantification in phytoplasma-infected plants. Bull. Insect. 64: S25-S26. **(IF: 0,592)**

Chapter 4
General Discussion

4.1 The complete nucleotide sequence of *Lupine mosaic virus*, LuMV - a new member in genus *Potyvirus*

Filamentous particles up to 950 nm in length are common characteristic for representative of family *Potyviridae* which is currently classified into six genera: *Ipomovirus, Macluravirus, Potyvirus, Rymovirus, Tritimovirus – with monopartite and Bymovirus,with bipartite genomes.*. Their genome is represented by a positive sense ssRNA; most of viruses are monopatite with the genome size 8 to 11 kbp, exception is genus *Bymovirus* with bipartite genome with 7.5 and 3.5 fragments; all viruses forming of cytoplasmic cylindrical and some of them nuclear inclusions. Except for some bymoviruses, they are transmissible by inoculation of sap; some potyviruses are carried of a small proportion of the seeds of some host species. All are vector transmitted by aphids, mites, whiteflies and by fungus *Polymyxa graminis* representatives of *Bymovirus.* Most of the viruses of this family restricted to host but some of them can infect wide range of plants (Shukla et al., 1998).

In this work was presented molecular characterization of a new potyvirus which was named *Lupine mosaic virus,* LuMV. It belongs to virus group which using RC strategy of translation: producing a large polyprotein precursor, which is then proteolytically cleavage by proteases of viral origin into 9-10 proteins. In addition to the similar genome size to potyviruses (8 kb to 11 kb), the predicted molecular mass was (364.3 kDa) for our LuMV and 10,113nt in length (Sarkisova et al., 2011). The putative open reading frame of the genomic RNA encodes a large polyprotein that is proteolytically processed into P1, helper component proteinase (HC-Pro), P3, cylindrical inclusion (CI) protein, 2 6K (6 kDa), viral genome-linked protein (VPg), nuclear inclusion proteins (NIa and NIb), and coat protein (CP), (Adams et al., 2005; Ng et al., 2006). In order to characterize the potyvirus complex it have been undertaken infected lupine samples, through the amplification, cloning and sequencing of its DNA. A reverse transcription followed by PCR with degenerate and/or specific primers let generate its complete sequence, which was submitted to GenBank database.

The fill-length DNA of LuMV from *L. polyphyllus* was obtained by PCR using a combination of random and specific primers (Gibbs et al., 1977) and was

characterized as a potyvirus belonging to the family *Potyviridae.* The sequence was 10,113 kbp in length a contained a large open reading frame (OF) that shown sequence similarity with several motives, which are conserved among viral CP, *rep*, and HC-Pro. In general, based on demarcation criteria for genus *Potyvirus* two viruses are different when the sequence identity at the CP aa level is the lower than 80% (Shukla et al., 1994; Adams et al., 2005; Desselberger et al., 2009). In this regard, *Lupine mosaic virus,* LuMV is confirmed to be a new member of the potyvirus group, since its highest sequence identity at the protein level to all previously described potyviruses was 60-64%. Both nucleotide and amino acid sequences of *Lupine mosaic virus*, LuMV further supported the close relationships between LuMV and other potyviruses. The capsid protein (CP) gene potentially coded for a protein of 30.3 kDa (Sarkisova et al., 2011). The DAG motive required for aphid transmission (Lopez-Moya et al., 1999) is presented in its N-terminus.

Plum pox virus, PPV and *Scallion mosaic virus,* ScaMV were identified by BLAST as closest relatives among the 63 completely sequenced potyviruses, sharing polyprotein aa sequence identities of 54 and 52%, respectively, with our *Lupine mosaic virus*, LuMV. Based on CP aa sequence comparisons, *Pepper veinal mottle virus*, PVMV, *Narcissus yellow stripe virus,* NYSV and *Chili veinal mottle virus*, ChiMoV were the closest relatives of Lupine *mosaic virus*, LuMV, with about 64% identity. Phylogenetic analysis was done to see evolutional relations more clearly and based on multiply alignment the greatest similarity (64%) was shown to *Lettuce mosaic potyvirus*, LMV by comparison of entire amino acids to 63 other completely sequenced potyviruses (Sarkisova et al., 2011).

Amongst all motives of the potyviral genome, the coat protein (CP) is the most frequently selected for studies of genetic diversity. The 3'-terminal position of the CP in the viral genome has contributed to the availability of a large number of CP sequences for different potyviruses (Shukla et al., 1994); comparison of CP aa sequences has allowed discrimination between species and strains (Shukla et al., 1988; Adams et al., 2005; Desselberger et al., 2009). An updated comparison of CP sequences has led to assign members of the family *Potyviridae* into genera, distinct species, related species, and strains (Ward et al., 1994; Shukla et al., 1994; Adams et al., 2005). Other regions of the potyviral genome, such as the 3' untranslated

region, have also been used to identify and classify potyviruses (Frenkel et al., 1989; Habera et al., 1994).

All nine motives (Ia to VIII) of the RNA-dependent RNA polymerases (Koonin et al, 1991) are present in the NIb protein of LuMV, but with the sequence SLMV2788 instead of TLMV in motive V.

The P1 protein is the most variable protein among potyviruses and functions as a serine protease for self-cleavage from HC-Pro (Adams et al., 2005; Verchot et al., 1991). The P1 protein is also known as silencing suppressor along with HC-Pro and has been suggested to play an important role in virus host range (Anandalakshmi et al., 1998; Salvador et al., 2008; Valli et al., 2006). P1 protein of LuMV with size of 51.3 kDa (Sarkisova et al., 2011) was in corresponding to other P1 proteins, which range in size from 29 kDa (for TVMV) to 63 kDa for *Papaya ringspot virus*, PRSV. P3 is known as membrane-bound protein and its functions are related to virus replication and movement (Cui et al., 2010; Eaitanaste et al., 2000).

Bothe HC-Pro and CP are known as multifunctional proteins that interact in multiple stages in the potyvirus infection cycle. The presence of PIPO motive in *Lupine mosaic virus,* LuMV was found to be similar, confirming that the LuMV is closely related to previously described potyviruses. Protease cleavage sites were recognized and were conservative as among other potyviruses. No putative recombination event was ascertained between *Lupine mosaic virus,* LuMV and the closely related viruses with the rdb v. 3.41 programs (results not shown). Phylogenetic analysis of the whole genome comparison showed that the position of LuMV polyprotein is close to that of *Lettuce mosaic virus,* LMV and *Plum pox virus,* PPV and *Panax virus Y*, PanVY (Fig. 13).

Lupine mosaic virus, LuMV very well meets the current species demarcation criteria in terms of genome sequence relatedness (less than 80% CP aa sequence identity and than 85% nt sequence identity over the whole genome) and, also due to the presence of different polyprotein cleavage sites (Adams et al, 2005; Berger et al, 2005; Desselberger et al., 2009). Therefore, *Lupine mosaic virus,* LuMV appears to be a new species in the genus *Potyvirus*. Therefore, this virus was assigned as a new potyvirus species to the genus, with name *Lupine mosaic virus,* LuMV.

4.2 Screening samples for plasmids presence including partial sequenced pBVK and pPEY

The only known proteins that have been found on phytoplasma plasmids to date are those involved in plasmid replication and maintenance, namely *rep, ssb* (single-strand binding protein), *DnaG* and *cop* (copy number protein), (Liefting et al., 2004; Nakashima et al., 1997; Nishigawa et al., 2003; Liefting et al., 2004). Genes on plasmids were shown recently are evolved rapidly compare to corresponding ones on phytoplasma chromosome. A good example, it is losing of their insect -vector transmissibility that was observed on clover plants generated by tissue culture over short period. It is caused by rearrangement of the plasmid DNA, which lack *ssb* protein gene and resulted the phytoplasma could no longer be transmitted by insects (Denes et al., 1992). The same observations were done by Nishigawa et al. (2002a), who determinated of lack *ssb* (single-strand DNA binding protein) gene and an uncharacterized putative membrane protein that exist in the wild-type line in a non-insect-transmissible line of onion yellows phytoplasma after 2 years of maintenance by plant grafting without insect vectors. For, sure plasmids could take a significant place in the adaptation to fast changing environmental conditions (Eberhard, 1990). Up to present, there are black spots in understanding of plasmids functions and still not much known but data increasing from year to year.

For the plasmid screening of 37 phytoplasma-infected samples were taken for the analysis. Gene -specific primers used for amplifying fragments in *rep* and *ssb* protein genes and later, two plasmids – pBVK and pPEY were subsequently taken for sequencing. Unfortunately, lack of sequence data has not given us to perform the full molecular characterization. Here is presented the short communication based on their partial sequences. Up to now, it can be conclude their presence in phytoplasmas, in spite of not complete understanding of their role in phytoplasma life cycle as a usual event. Using gene specific primers for conservative genes screening suspected phytoplasma samples was performed. DNA isolation from 0, 1-0, 2 mg of fresh material was following by the PCR cycle for amplifying short fragments universal primers of our design (Petrzik et al., 2011) in conservative *rep* gene showed good results in testing.

Rep genes were amplified by universal primers' pairs: 426 K8/K9. Additionally, several pairs were designed for two selected phytoplasma plasmids - pBVK and pPEY that were chosen to continue sequencing. The conservative motives were recognized in pPEY in replicase gene (Sarkisova et al., 2012).A presence of these motives is pointed on the structure of replicase gene which is shared by plasmids pLS1 family (Oshima et al., 2001) which is known to be replicated by RC mechanism (Solar et al., 1993; Khan, 1997). A neighbor-joining (NJ) tree for the amino acid sequences of plasmids family pLS1 including pBVK and pPEY was created and revealed that plasmid pKMK1 from *Mycoplasma mycoides* is the closest and sharing one cluster (Fig. 16). The interesting property of some plasmids belonging to this family is a presence a helicase domains of eukaryotic origin in C-terminal part, which was as firstly, described for the plasmid from phytoplasma *Onion yellows*, pOYW by Oshima et al., (2001). Unfortunately, there is a lack data of complete C-terminal part of *rep* gene, however, the highest % of similarities which our two plasmids pBVK and pPEY showed to pOYW give us to assume the similar genome organization and might be keeping the same type of plasmid. A phylogenetic analysis was done in order to see the position of our two plasmids toward of the *rep* of pOYW which having helicase domain. Creating phylogenetic tree was based of alignment avaible parts of *rep* gene and was found closer relations to analogous from the plant geminiviruses and less eukaryotic viruses. These plasmids showed high similarity to an extrachromosomal DNA from *Onion yellows* phytoplasma (Fig.17). It was clear that pPEY and pBVK are more closely related to geminiviruses :*Beet curly top virus*, BCTV*, Bean golden yellow mosaic virus, BGMV* and TLCV, *Tomato leaf curl virus* what is corresponding with results for pOYW and sub cluster with a pKMK1 and less to non-plant eukaryotic viruses – PCV, *Porcine circovirus,* BFDV, *Beak* and *feather disease virus* (Fig.17).

Up to present, both plasmids have been partially sequenced; pBVK and pPEY are approximately 1300 - 1500 nt in length. The sequenced region was spanned: not completed putative *rep gene,* including full-length sequence of *ssb* and completely/partially two unknown proteins for pBVK / pPEY, respectively (Sarkisova et al., 2012). The nucleotide sequence derived from these partial sequences of both plasmids showed homology with functionally unknown, *ssb* proteins encoded by phytoplasma chromosome of *Brassica napus*, *Rehmania*

glutinosa and *Aster yellows witches'-broom* phytoplasma plasmids. The deduced nucleotide sequence of pBVK derived from *rep* gene shared high similarity- 97% with hypothetical replication genes protein encoded on the plasmids from onion yellows and *Rehmania glutinosa* agent phytoplasmas whereas pPEY to *Brassica napus* phytoplasma and *Onion yellows* phytoplasmas plasmid.

The obtained results showed the plasmid' presence in more than half of infected samples; their phylogenetic relatedness was not in correspondence to relatedness between of phytoplasmas in groups and subgroups (Sarkisova et al., 2012). Continuing work with other samples which are not discussed here let to do a conclusion that there is type of plasmid which carry *Onion yellows* phytoplasma, f.i, looks like an usual event for members *Aster' yellows* group and more probably, in other strains based on data our two partially sequenced plasmids.

A variety of supporting assays have been used as p29 polymerase amplification, polymerase chain reaction (PCR), selective amplification of phytoplasma plasmid DNA using universal and gene-specific primers.

Conclusions

- *Lupine mosaic virus*, LuMV – a new member in genus *Potyvirus* with the closest phylogenetic relatedness to *Lettuce mosaic virus, LMV*
- The complete nucleotide sequence of the LuMV ssRNA was determined and submitted to GenBank under accession number HM748648
- CP, *rep* motives and the whole genome sequence, were used for the phylogenetic analysis and shown - *Lupine mosaic virus*, LuMV meets well the current species demarcation criteria for genus *Potyvirus*
- In agreement with VIIIth report of the ICTV the proposal to use CP to the taxonomic analysis, (less than 80% CP aa sequence identity and than 85% nt sequence identity over the whole genome) aa identities as species demarcating criteria
- Results of the phytoplasma plasmids screening revealed their presence as an usual event for phytoplasmas, those with the genome organization similar to pOYW – a type of plasmid which occur widely
- *Rep* gene structure let conclude belonging of both plasmids pBVK and pPEY to prokaryotic plasmids of pLS1 family which are commonly replicate via a rolling circle replication (RCR) strategy
- A maximum likelihood tree for the amino acid sequences of the N-terminal region of the *rep* protein of detected plasmids showing the phylogenetic relations which are not always corresponding to those between phytoplasmas which carry out these plasmids

References

1. Adams, M.J, Antoniw, J.F, Fauquet C.M. 2005. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch. Virol. 150: 459–479.
2. Aleman, M.E, Marcos J.F, Brugidou C, Beachy R.N, Fauquet C. 1996.The complete nucleotide sequence of yam mosaic virus (Ivory Coast isolate) genomic RNA. Arch. Virol. 141: 1259-1278.
3. Allison, R. F., Janda, M., Ahlquist, P. 1989. Sequence of cowpea chlorotic mottle virus RNAs 2 and 3 and evidence of a recombination event during bromovirus evolution. Virology 172: 321-330.
4. Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., Vance, V. B. 1998. A viral suppressor of gene silencing in plants. Proc. Natl. Acad. Sci. 95: 13079-13084.
5. Atreya, C.D. 1992. Application of genome sequence information in potyvirus taxonomy: an overview. Arch. Virol. Suppl. 5: 17-23.
6. Atreya, C.D., Atreya, P.L., Thornbury, D.W., Pirone, T.P. 1992. Site- directed mutations in the potyvirus HC-Pro gene affect helper component activiry, virus accumulation, and symptom expression in infected tobacco plants. Virology 191: 106-111.
7. Barnett, O.W.1991.Potyviridae, a proposed family of plant viruses. Arch.Virol.118: 139-141.
8. Bariana, H.S., Shannon, S.L., Chu, P.W.G., Waterhouse, P.M.1994. Detection of five seed-borne legume viruses in one sensitive multiplex polymerase chain reaction test. Phytopathology 84: 1201-1205.
9. Becker, Y. 2000. Evolution of viruses by acquisition of cellular RNA or DNA nucleotide sequences and genes: An introduction. Virus Genes 21: 7-12.
10. Berger, P.H., Adams, M.J., Barnett, O.W., Brunt, A.A., Hammond ,J., Hill J.H., Jordan R.L., Kashiwazaki, S., Rybicki, E., Spence,N., Stenger, D.C., Ohki S.T., Uyeda, I., van Zaayen, A., Valkonen, J., Vetten, H.J. 2005. *Potyviridae*. In: Fauquet, C.M., Mayo, M.A., Maniloff, J.
11. Bertaccini, A. 2007. Phytoplasmas: diversity, taxonomy, and epidemiology. Front. Biosci. 12: 673-689.
12. Blanc, S., Ammar, E. D., Garcia-Lampasona, S., Dolja, V. V., Llave, C., Baker, J., and Pirone, T. P. 1998. Mutation in potyvirus helper component protein: effects on interactions with virions and aphid stylet. J. Gen. Virol. 79: 3119-3122.
13. Bos, L. D., Hagedorn J., Quatz, L. 1960**.** Suggested procedures for international identification of legume viruses. Eur. J. Plant Pathol. 66: 328-343.
14. Brown, J.K., 1990. An update on the whitefly-transmitted geminiviruses in the Americas and the Caribbean Basin. FAO Plant Prot. Bull. 39: 5–23.

15. Brunt A., Crabtree K., Gibbs A.1990. In: Viruses of Tropical Plants. 707. C.A.B, International, Wallingford, UK.
16. Bubanovic I., Najman S., Andjelkovic Z. 2005. Origin and evolution of viruses: Escaped DNA/RNA sequences as evolutionary accelerators and natural biological weapons. Med. Hypotheses 65: 868–872.
17. Carrington J.C., Cary S.M., Parks T.D, Dougherty W.G. 1989a. A second proteinase encoded by a plant potyvirus genome. EMBO J. 8: 365-370.
18. Chopra, I., Roberts, M. 2001. Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microb. Mol. Biol. Rev. 65: 232–260.
19. Christensen, N. M., Axelsen, K. B., Nicolaisen, M., Schulz, A. 2005. Phytoplasmas and their interactions with hosts. Trends Plant Sci. 10: 526–535.
20. Chung, B.Y.M., Miller, W.A., Atkins, J., Firth, A.E. 2008. An overlapping essential gene in the Potyviridae. Proc. Nat. Acad. Sci. 105:5897–5902.
21. Clark, M.F., Adams, A.N. 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. J. Gen. Virol. 34: 475–483.
22. Cotton, S., Grangeon, R., Thivierge, K., Mathieu, I., Ide, C., Wei, T., Wang, A., Laliberte, J. F. 2009. Turnip mosaic virus RNA replication complex vesicles are mobile, align with microfilaments, and are each derived from a single viral genome. J. Virol.83: 10460–10471.
23. Cronin S., Verchot J., Haldeman-Cahill, R., Scaad, M.C., Carrington, J.C.1995. Long distance movement factor: a new novel transport function of the potyvirus helper-component-proteinase. Plant Cell 7: 549-559
24. Cui, X., Wei, T., Chowda-Reddy, R. V., Sun, G., Wang, A. 2010. The Tobacco etch virus P3 protein forms mobile inclusion via the early secretory pathway and traffics along actin microfilaments. Virology 397: 56-63.
25. Davies, J.E, Rownd, R.1972. Transmissible multiple drug resistance in Enterobacteriaceae. Science 176: 758–768.
26. Denes, A. S., Sinha, R. C. 1992. Alteration of clover phyllody mycoplasma DNA after in vitro culturing of phyllody-diseased clover. Can. J. Plant Pathol. 14: 189–196.
27. Desselberger, U., Ball, L.A. (Eds). 2009. Virus taxonomy. VIIIth report of the international committee on taxonomy of viruses. Elsevier/ Academic Press 819–829.
28. Desjardins, C., Eisen, J.A., Nene V. 2005. New evolutionary frontiers from unusual virus genomes. Gen. Biol. 6: 212–3.
29. Doi, Y., Teranaka, M., Yora, K., Asuyama, H., 1967. Mycoplasma-or PLT group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches' broom, aster yellows, or paulownia witches' broom. Ann. Phytopath. Soc. Japan 33: 259-266.

30. Domier, L. L., Shaw, J. G., Rhoads, R. E. 1987. Potyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. Virology 158: 20-27.
31. Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J. Novella, I. S., Holland, J. J. 1996. Basic concepts in RNA virus evolution. FASEB J. 10: 859-864.
32. Domingo, E., Holland, J. J. 1997. RNA virus mutations and fitness for survival. Ann. Rev. Microbiol. 51: 151-178.
33. Dougherty, W.G., Carrington, J.C .1989. Expression and function of potyviral gene products. Ann. Rev. Phytopath. 26: 123-143.
34. Eaitanaste, S., Juricek, M., Yap, Y. 2007. C-terminal hydrophobic region leads PRSV P3 to endoplasmic reticulum. Virus Genes 35: 611-617.
35. Eberhard, W. G. (1990). Evolution in bacterial plasmids and levels of selection. Quarter. Rev. Biol. 65: 3–22.
36. EPPO quarantine pest.1996. Data Sheets on Quarantine Pests. Bean golden mosaic bigeminivirus. Prepared by CABI and EPPO for the EU under Contract 90/399003**.**
37. Fauquet, C. 2002. Geminivirus species demarcation criteria study case. Webpage Geminiviridae: www.danforthcenter.org/iltab/geminiviridae/speciesdemarcation.htm
38. Fauquet, C.M., Bisaro, D.M., Briddon, R.W., Brown, J.K., Harrison, B.D., Rybicki, E.P., Stenger, D.C., Stanley, J .2003. Revision of taxonomic criteria for species demarcation in the family *Geminiviridae*, and an updated list of begomovirus species. Arch. Virol. 148: 405–421.
39. Frenkel, M.J., Ward, C.W., Shukla, D.D. 1989. The use of 3- non-coding nucleotide sequences in the taxonomy of potyviruses: application to *watermelon mosaic virus* 2 and *soybean mosaic virus* – N. J. Gen Virol. 70: 2775-2783.
40. Galie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C., Wilson, T.M.A. 1987. A comparison of eukaryotic viral 5- leader sequences as enhancers of mRNA expression *in vivo.* Nucl. Acids Res. 15: 8693-8711.
41. Gal-On, A. 2000.A point mutation in the FRNK motive of the potyvirus helper component-protease gene alters symptom expression in cucurbits and elicits protection against the severe homologous virus. Phytopathology 90: 467–473.
42. García-Arenal F., Fraile A., Roossinck M.J. 2008. Questions and Concepts in Plant Virus Evolution: a Historical Perspective. Plant Virus Evol.1-14.
43. Christensen, N. M., Axelsen, K. B., Nicolaisen, M., Schulz, A. 2005. Phytoplasmas and their interactions with hosts. Trends Plant Sci. 10: 526–535.
44. Gibbs, A., Mackenzie, A. 1977. A primer pair for amplifying part of the genome of allpotyvirids by RT-PCR. J.Virol. Methods 63: 9-16

45. Gibbs, A.J., Mackenzie, A.M., Gibbs, M.J. 2003. The 'potyvirus primers' will probably provide phylogenetically informative DNA fragments from all species of *Potyviridae*. J. Virol. Methods 112: 41–44.
46. Gorbalenya, A.E., Koonin E. V. 1993. Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin.Struct. Biol. 3:419-429.
47. Gorbalenya, A. E. 1995. Origin of RNA viral genomes; approaching the problem by comparative sequence analysis. In: Molecular basis of virus evolution. (Eds.) by Cambridge University Press, Cambridge. 49-65.
48. Ha, C., Coombs, S., Revill, P.A., Harding, R.M., Vu, M., Dale, J.L.2008. Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Arch .Virol 153:25–36.
49. Habera, L. F., Berger, P. H., Reddick, B. B. 1994. Molecular evidence from 3'-terminus sequence analysis that tobacco vein-banding mosaic virus is a distinct member of the potyvirus group. Arch. Virol.138: 27-38.
50. Hamilton, I.R.1997.Legume virus research in Canada: a retrospective and a view of the future. Can.J. Plant Pathol. 19: 208-214.
51. Holland, J., Domingo, E. 1998. Origin and evolution of viruses. Virus Genes 16: 13-21.
52. Huet, H., Gal-On, A., Meir, E., Lecoq, H., Raccah, B. 1994. Mutations in the helper component protease gene of zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility. J. Gen. Virol. 75: 1407-1414.
53. Johansen, E., Rasmussen, O.F., Heide, M., Borkhardt, B. 1991. The nucleotide sequence of pea seed-borne mosaic virus RNA. J. Gen. Virol. 72: 2625-2632
54. Kasschau, K. D., Carrington, J. C. 2001. Long-distance movement and replication maintenance functions correlate with silencing suppression activity of potyviral HC-Pro. Virology 285:71-81.
55. Kay, D.E, 1979. Food legumes. Trop. Devel. and Res. Inst. London. Field Crop Abst. 26: 157-160.
56. Khan, S. A. 1997. Rolling-Circle Replication of Bacterial Plasmids. Microb. Mol. Boil. Revs. 61: 442–455.
57. Kirkpatrick, B. C. 1992. Mycoplasma-like organisms: plant and invertebrate pathogens. 4050- 4067. In: The Prokaryotes. (2nd ed) Balows, A., Truper, H.G., Dworkin M., Harder W., Schleifer, K.H. (Eds.) by Springer-Verlag, New York.
58. Kobe, B., Deisenhofer, J. 1994. Proteins with leucine – rich repeats. Curr. Opin. Struc. Biol. 5: 409 -416
59. Koonin, E.V., Choi, G.H., Nuss, D.L., Shapira, R., Carrington, J.C.1991. Evidence for common ancestry of a chestnut blight hypovirulence-

associated double-stranded RNA and a group of positive strand RNA plant viruses. Proc Nat. Acad. Sci. 88: 10647–10651.

60. Koonin, E. V., Ilyina, T. V. 1992. Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. J. Gen. Virol. 73: 2763-2766.
61. Kuboyama, T., Huang, C. - C., Lu, X., Sawayanagi, T., Kanazawa, T., Kagami, T., Matsuda, I., Tsuchizaki, T., Namba, S. 1998. A plasmid isolated from phytopathogenic onion yellows phytoplasma and its heterogeneity in the pathogenic phytoplasma mutant. Mol. Plant-Microbe Interact. 11:1031-1037.
62. Kumari, S.G., Makkouk, K.M. 2007. Virus diseases of faba bean (*Vicia faba* L.) in Asia and Africa. Plant Viruses 1: 93–105.
63. Kuske, C. R., Kirkpatrick, B. C. 1990. Identification and characterization of plasmids from the western aster yellows mycoplasma like organism. J. Bacteriol. 172:1628-1633.
64. Lain, S., Reichmann, J.L., Garsia, J.A. 1989. The complete nucleotide sequence of plum pox potyvirus RNA. Virus Res. 13: 157-172.
65. Lain, S., Reichmann, J.L., Garcia, J.A.1990. RNA helicase: a novel associated with a protein encoded by a positive strand RNA virus. Nucl. Acids Res. 18: 7003-7006.
66. Lee, I-M., Gundersen-Rindal, D.E., Davis, R.E., Bartoszyk, I.M. 1998. Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences. Int. J. Syst. Bacteriol. 48: 1153-1169.
67. Lee, I. M., Davis, R. E., Gundersen-Rindal, D. E. 2000. Phytoplasma: phytopathogenic Mollicutes. Ann. Rev. Microbiol. 54: 221–255.
68. Lin C-L., Zhou T., Li H-F., Fan, Z-F., Li, Y., Piao, C. –G., Tian, G. -Z. 2009. Molecular characterisation of two plasmids from paulownia witches'-broom phytoplasma and detection of a plasmid encoded protein in infected plants. Eur. J. Plant Pathol. 123: 321–330.
69. Liefting, L. W., Shaw M.E., Kirkpatrick B.C. 2004. Sequence analysis of two plasmids from the phytoplasma beet leafhopper-transmitted virescence agent. Microbiol. 150: 1809–1817.
70. Liefting, L.W., Andersen M.T., Lough T.J., Beever R.E. 2006. Comparative analysis of the plasmids from two isolates of "Candidatus Phytoplasma australiense. " Plasmid 56: 138–144.
71. Loebenstein, G., Thottappilly, G. 2004. Virus and virus- like diseases of major crops in developing countries. (Eds.) by Kluwer Academic Publishers, Doordrecht/ Boston/ London.442
72. Lopez-Moya, J.J., Wang, R.Y., Pirone, T.P. 1999. Context of the chat protein DAG motive affects potyvirus transmissibility by aphids. J. Gen. Virol. 80: 3281–3288

73. Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., Scheele, G.A. 1987. Selection of AUG initiation codons differs in plants and animals. EMBO J. 6: 43–48.
74. Makkouk, K.M., Comeau A. 1994. Evaluation of various methods for the detection of barley yellow dwarf luteovirus by the tissue-blot immunoassay and its use for BYDV detection in cereals inoculated at different growth stages. Eur. J. Plant Pathol. 100: 71–80.
75. Makkouk, K.M., Kumari, S.G., 1996. Detection of 10 viruses by the tissue-blot immunoassay (TBIA). Arab. J. Plant Protect. 14: 3–9.
76. Makkouk, K.M., Kumari S.G. 2009. Epidemiology and integrated management of persistently transmitted aphid-borne viruses of legume and cereal crops in West Asia and North Africa. Virus Res. 141: 209–218.
77. Mali, V.R., Thottappilly, G.1986. Virus diseases of cowpea in the tropics. In: Raychandhuri, S.P., Varma, J.P. Rev. Tropical Plant Dis. 3: 361–403.
78. Maniloff, J. 1988. Mycoplasma viruses. Crit. Rev. Microbiol. 15: 339-389.
79. Mansour A., Al-Musa, A., Vetten, H.J., Lesemann D.E. 2003. Properties of a Cowpea Mild Mottle Virus (CPMMV) Isolate from Eggplant in Jordan and Evidence for Biological and Serological Differences Between CPMMV Isolates from Leguminous and Solanaceous Hosts. J. Phytopath. 128: 539-547.
80. Margulis, L., Sagan, D. 1997. Microcosmos: four billion years of evolution from our microbial ancestors. University of California Press.
81. Margulis, L., Sagan D. What is life? University of California Press. 2000.
82. Matthewes, R.E.F.1991. Plant Virology. (3d Ed.) by Academic Press, San Diego.
83. McGeoch, D. J., Davison A. J. 1999. The molecular evolutionary history of the herpes viruses. 441 -465. In: Origins and Evolution of Viruses by Domingo, E., Webster, R., and Holland, J. (Eds.) by Acad. Press, London.
84. Morales, F.J., Castano, M., Arroyava, J.A., Ospina, M.D., Calvert, L.A. 1995. A sobemovirus hindering the utilization of Calopogonium mucunoides as a forage legume in the lowland tropics. Plant Dis. 79: 220-1224.
85. Morales, F.J., Niessen, A.I., Ramirez, B., Castano, M. 1990. Isolation and partialcharacterization of a geminivirus causing bean dwarf mosaic. Phytopathology 80: 96-101.
86. Nakashima, K., Hayashi, T. 1997. Sequence analysis of extrachromosomal DNA of sugarcane white leaf phytoplasma. Ann. Phytopath. Soc. Jap. 63: 21–25.
87. Nicolas, O., Laliberte, J.F. 1992. The complete sequence of turnip mosaic potyvirus RNA. J. Gen. Virol. 73: 2785-2793.
88. Nishigawa, H., Miyata, S-I., Oshima, K., Sawayanagi, T., Komoto, A., Kuboyama, T., Matsuda, I.,Tsuchizaki, T.,Namba, S. 2001. In planta expression of a protein encoded by the extrachromosomal DNA of a

phytoplasma and related to geminivirus replication proteins. Microbiology 147: 507–513.

89. Nishigawa, H., Oshima, K., Miyata, S., Ugaki, M., Namba, S.2003. Complete set of extrachromosomal DNAs from three pathogenic lines of onion yellows phytoplasma and use of PCR to differentiate each line. J. Gen. Plant Pathol. 69: 194–198.
90. Novichkov, P. S., Omelchenko, M. V., Gelfand, M. S., Mironov, A. A., Wolf, Y. I., Koonin, E. V. 2004. Genome-wide molecular clock and horizontal gene transfer in bacterial evolution. J. Bacteriol. 186: 6575–6585.
91. Oh, C.S., Carrington, J.C. 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. Virology 173: 692-699.
92. Oshima K., Kakizawa S.,Nishigawa H., Kuboya T., Miyata S-I., Ugaki M., Namba S.. 2001. A Plasmid of Phytoplasma Encodes a Unique Replication Protein Having Both Plasmid- and Virus-like Domains: Clue to Viral Ancestry or Result of Virus/Plasmid Recombination? Virology 285: 270-277
93. Oshima, K., Shiomi, T., Kuboyama, T., Sawayanagi, T., Nishigawa, H., Kakizawa, S., Miyata, S., Ugaki, M., Namba, S. 2001. Isolation and characterization of derivative lines of the onion yellows phytoplasma that do not cause stunting or phloem hyperplasia. Phytopathology 91: 1024–1029.
94. Paliwal.Y.C.1982. Virus diseases of alfalfa and biology of alfalfa mosaic virus in Ontario and Quebec. Can. J. Plant Pathol. 4: 175-179.
95. Peng, Y. H., Kadoury, D., Gal-On, A., Huet, H., Wang, Y., Raccah, B. 1998. Mutations in HC-Pro gene of *Zucchini yellow mosaic potyvirus*: effects on aphid transmission and binding to purified virions. J. Gen. Virol. 79: 897-904.
96. Panopoulos, N., Peet, R. C. 1985. The molecular genetics of plant pathogenic bacteria and their plasmids. Ann. Rev. Phytopath. 23: 381-419.
97. Petrzik, K., Krawczyk, K., Zwolinska, A. 2011.Two high-copy plasmids found in plants associated with strains of ‘‘Candidatus Phytoplasma steris’’.Plasmid 66: 122–127.
98. Puurand, U., Makinen, K., Paulin, L., Saarma, M. 1994. The nucleotide sequence of potato virus A genomic RNA and its sequence similarities with other potyviruses. J Gen. Virol. 75: 457-461.
99. Puustinen, P., Makinen, K. 2004. Uridylylation of the potyvirus VPg by the viral replicase NIb correlates with the nucleotide binding capacity of VPg. J. Biol. Chem. 279: 38103–38110.
100. Rankin, W. H., Fraser, W.P. 1922. Survey of prevalence of common plant diseases in the Dominion in Canada, 1921. Canade Expt. Farms. Div.Bot.Ann.Rep. Canad. Plant Dis. Survey 2:1-62.
101. Rao, A.L.N., C.Huriki. 1985a. Occurrence and identification of Lucerne transient streak virus in Alberta, Canada. Plant Dis. 69: 610-612.

102. Rao, A.L.N., Huriki C. 1985b. Clover primary leaf necrosis virus, a strain of red clover necrotic mottle virus. Plant Dis. 69: 959-961.
103. Rao, A.L.N., Lakshman, D.K., Ohki, S.T., Huriki, C. 1987. Identification of a new serotype and antigenic relationships among six strains of red clover necrotic mosaic virus. Phytopathology 77: 995-999.
104. Razin, S. 1985. Molecular biology and genetics of mycoplasma (Mollicutes). Microbiol. Rev. 49: 419-455.
105. Rodriduez-Cerezo, E., Gamble-Klein P., Shaw J.G. 1991. A determinant of disease symptom severity is located in the 3'- terminal noncoding region of the RNA of a plant virus. Proc. Natl. Acad. Sci. 88: 9863-9867.
106. Rossinck, M. J. 1997. Mechanisms of plant virus evolution. Ann. Rev. Phytopath. 35: 191-209.
107. Roudet-Tavert, G., Michon, T., Walter, J., Delaunay, T., Redondo, E., Le Gall, O. 2007. Central domain of a potyvirus VPg is involved in the interaction with the host translation initiation factor eIF4E and the viral protein HcPro. J. Gen. Virol. 88: 1029–1033.
108. Saldarelli, P., Rowhani, A., Routh, G., Minafra, A., Digiaro, M.1998. Use of degenerate primers in a RT-PCR assay for the identification and analysis of some filamentous viruses, with special reference to clostero- and vitiviruses of the Grapevine. Eur. J. Plant. Pathol. 104: 945-950.
109. Saleh, N., Horn, N.M., Reddy, D.V.R., Middleton, K.J.2005. Peanut stripe virus in Indonesia. Eur. J. Plant. Pathol. 95:123-127.
110. Salvador, B., Saenz, P., Yanguez, E., Quiot, J. B., Quiot, L., Delgadillo, M. O., Garcia, J. A., Simon-Mateo, C. 2008. Host-specific effect of P1 exchange between two potyviruses. Mol. Plant Pathol. 9: 147-155.
111. Sanjuan, R., Nebot, M.R., Chirico, N., Mansky L.M., Belshaw, R. 2010. Viral mutation rates. J. Virol. 84: 9733–9748.
112. Schneider, B., Maurer, R., Saillard, C., Kirkpatrick, B., Seemüller, E. 1992. Occurrence and relatedness of extrachromosomal DNAs in plant pathogenic mycoplasma like organisms. Mol. Plant-Microbe Interact. 5: 489–495.
113. Seemüller, E., Marcone, C., Lauer, U., Ragozzino, A., Goschl, M., 1998. Current status of molecular classification of the phytoplasmas. J. Plant Pathol. 80: 3–26.
114. Stuart, G., Moffett, K., Bozarth, R. F. 2004. A whole genome perspective on the phylogeny of the plant virus family *Tombusviridae.* Arch.Virol. 149: 1595-1610.
115. Shukla, D. D., Ward, C. W. 1988. Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. J. Gen. Virol. 69: 2703-2710.
116. Shukla, D. D., Frenkel, M. J., Ward, C. W. 1991. Structure and function of the potyvirus genome with special reference to the coat protein-coding region. Can. J. Plant Pathol. 13: 178-191.

117. Shukla, D.D, Ward, C.W., Brunt, A.A. 1994. The Potyviridae. CAB International, Wallingford, UK., pp. 516.
118. Shukla, D.D., Ward, C.W. 1989. Identification and classification potyviruses on the base of capsid protein sequence data and serology. Arch. Virol. 106: 171-200.
119. Shukla, D. D., Ward, C. W., Brunt, A. A., Berger, P. H. 1998. Description of plant viruses:Potyviridae.Webpage: http://www.dpvweb.net/dpv/showdpv.php?dpvno=366/
120. Simon, A. E., Bujarski, J. J.1994. RNA-RNA recombination and evolution in virus-infected plants. Ann. Rev. Phytopath. 32: 337-362.
121. Solar G., Moscoso, M., Espinosa M. 1993. Micro Review Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: a wall falls. Mol. Microbiol. 8: 789-796.
122. Taiwo, M.A. 2003. Viruses infecting legumes in Nigeria: case history. In: Hughes J'd' A., Odu B. (Eds.) by Plant Virol. in Sub-Saharan Africa. Proceed. Plant Virol. IITA, Ibadan, Nigeria. 364–378.
123. Tamura K., Dudley, J., Nei M., Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Bio. Evol. 24: 1596–1599.
124. The Angiosperm Phylogeny Group, 1998. "An ordinal classification for the families of flowering plants". Ann. of the Missouri Bot. Garden. 85: 531–553.
125. Thottappilly, G., Rossel, H.W.1985. World-wide occurrence and distribution of virus diseases. 155-171. In: Cowpea research, production and utilization. (Eds.) by Singh, S. R., Rachie, K. O. John Wiley and Sons, Chichester, UK.
126. Thottappilly, Y. G., Rossel, H.W.1992. Virus diseases of cowpea in tropical Africa. Trop. Pest Management 38: 337–348.
127. Urcuqui-Inchima, S., Haenni, A-L., Bernardi, F. 2001. Potyvirus proteins: a wealth of functions. Virus Res. 74: 157–175.
128. Valli, A., Martín-Hernández, A. M., López-Moya, J. J., García, J. A. 2006. RNA silencing suppression by a second copy of the P1 serine protease of *Cucumber vein yellowing ipomovirus* (CVYV), a member of the family *Potyviridae* that lacks the cysteine protease HCPro. J. Virol. 80: 10055–10063.
129. van Regenmortel, M.H.V., Bishop, D.H.L., Fauquet, C.M., Mayo, M.A., Maniloff, J., Calisher, C.H. 1997. Guidelines to the demarcation of virus species. Arch. Virol. 142: 1505–1518.
130. van Regenmortel, M.H.V.2000.Introduction to the species concept in virus taxonomy. In Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. (Ed) by van Regenmortel, M.H.V., Fauquet, C., Bishop, D.H.L, Carstens, E.B., Estes, M.K.,

Lemon, S.M., McGeogh, D.J., Maniloff, J., Mayo, M.A., Pringle, C.R.,Wickner, R.B. San Diego. Academic Press, pp. 3-16.

131. Varrelmann, M., Maiss, E., Pilot, R., and Palkovics, L. 2007. Use of pentapeptide-insertion scanning mutagenesis for functional mapping of plum pox virus helper component proteinase suppressor of gene silencing. J. Gen. Virology 88: 1005-1015.

132. Verchot, J., Koonin, E.V., Carrington, J.C. 1991.The 36- kDa protein from the N-terminal of the potyviral polyprotein functions as a third-encoded proteinase. Virology 185: 527-535.

133. Verchot, J, Herndon, K.L, Carrington, J.C. 1992. Mutational analyses of the tobacco etch potyviral 35kDa proteinase: identification of essential residues and requirements for autoproteolysis. Virology 190: 298 -306.

134. Vivian, A., Murillo, J., Jackson, R. W. 2001. The roles of plasmids in phytopathogenic bacteria: Mobile arsenals? Microbiology 147: 763–780.

135. Ward, C. W., Shukla, D. D. 1994. Structure and variation of potyviruses. In Virol. in the Tropics. 43-61. (Ed.) by Rishi N., Ahuja, K.L., Singh B.P. New Delhi: Malhotra Publishing House.

136. Wei, T., Zhang, C., Hong, J., Xiong, R., Kasschau, K. D., Zhou, X. P., Carrington, J. C., and Wang, A. M. 2010a. Formation of complexes at plasmodesmata for potyvirus intercellular movement is mediated by the viral protein P3N-PIPO. PLoS Pathogens 6: e1000962.

137. Wegrzyn, G. 2005. What does “plasmid biology” currently mean? Summary of the Plasmid Biol. 2004 meeting. Plasmid 53: 14–22.

138. Wen, R.H., Hajimorad, M.R. 2010. Mutational analysis of the putative pipo of soybean mosaic virus suggests disruption of *pipo* protein impedes movement. Virology 400:1-7.

139. Winans, S. C. 1992. Two-way chemical signaling in *Agrobacterium*-plant interactions. Microbiol. Rev. 56**:** 12–31.

140. Worobey, M., Rambaut, A., Pybus, O. G., Robertson, D. L. 2002. Questioning the evidence for genetic recombination in the 1918 "Spanish flu" virus. Science 296: 211.

141. Zuker, M., Mathews, D. H., Turner, D. H. 1999. Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide In *RNA Biochemistry and Biotechnology*, J. Barciszewski & B.F.C. Clark. (Eds.) by NATO ASI Series, Kluwer Academic Publishers.