Szent István University
Doctoral School of Biological Sciences

Molecular analysis of the coat protein genes of maize dwarf mosaic virus (MDMV) populations

Main points of the thesis

Gyöngyvér Mónika Gell

Supervisors: Dr Ervin Balázs, ARI HAS
Dr László Hornok, SIU

Gödöllő – Martonvásár
2011

1. BACKGROUND AND AIMS

1.1 Background

The production of maize has a long history in Hungary. First brought to Europe by Spanish traders, it reached Hungary via Turkey and has played a decisive role in the country's agriculture ever since. Cereals make up the larger part of crop production in Hungary (50–70%), with maize (25–30%) and wheat (18–30%) the dominant crops on cereal-growing areas. Maize occupies the largest area of all field crops, so its role in agriculture is second to none.

Maize grain yields are greatly influenced by quantitative and qualitative losses caused by pests and pathogens. One such pathogen, the most important virus disease attacking maize, is the maize dwarf mosaic virus (MDMV).

This virus, which belongs to the *Potyviridae* family and is one of the most significant viral pathogens of monocotyledonous plants, was first recorded in the United States in the 1960s (Janson and Ellett 1963, Williams and Alexander 1965), but was soon detected in Hungary too (Szirmai and Paizsné 1963). Its host plants are maize, johnson grass, sugarcane, millet, sorghum species and other grass species (Williams and Alexander 1965). MDMV can be spread with seed (0.007–0.4%; Hill et al. 1974, Midel et al. 1984), mechanically, or by leaf aphids in a non-persistent manner (Toler 1985).

Different maize varieties exhibit various levels of sensitivity to virus infection. The most frequent symptoms are mosaicism, dwarfing, poorly developed tassels and poor seed setting (Revers et al. 1999). The virus infection causes serious damage to agriculture (Oertel et al. 1997) with yield losses of as much as 42% (Peti 1983, Sum et al. 1979, Szirmai 1968). Considering the fact that the level of infection for specific varieties may be as high as 80% (Tóbiás et al. 2003), the damage caused by MDMV is of great economic significance. The economic damage consists of the high proportion of ears with aborted ends, a reduction in ear size and a decrease in the thousand-kernel weight (Toldiné Tóth 2008).

While other members of the SCMV (*Sugarcane mosaic virus*) subgroup, which are serologically related to MDMV, have been analysed in detail by scientists investigating the viruses infecting monocotyledons, very little attention has been paid to the importance of MDMV. This is confirmed by gene bank data, which contain only a limited number of MDMV sequences. As MDMV is the dominant virus in the subgroup in Hungary, there is every justification for a wider analysis of the virus.

The variability of the genetic material of MDMV and the possible pathological consequences of this raised the need for a more detailed analysis of the symptom determinants and populations of

the virus. In experiments on pathogen-derived resistance (PDR) the coat protein genes of the viruses, or segments of these genes, are widely used to induce virus resistance. The coat protein is multifunctional, being an important symptom and host specificity determinant, while also being essential for aphid transmission and for the movement of the virus both from cell to cell and over longer distances. As it has such a complex role, the determination of its primary structure is a good basis for the genetic analysis of the virus.

1.2 Aims

Due to the properties of the coat protein outlined above, the present work involved comparative analysis of the coding region of the coat protein (CP).

The work included the following studies:

1. The determination of the genetic background of maize dwarf mosaic virus (MDMV) based on the primary structure of the coat protein gene of isolates collected from two maize growing and breeding areas in Hungary, in order to map the composition and stability of the Hungarian virus population.

2. The preparation of a construct containing the MDMV coat protein gene, capable of inducing virus resistance in transgenic plants. Elaboration of suitable transformation methods.

3. The development of a full-length, hopefully virulent MDMV clone.

2. MATERIALS AND METHODS

2.1 Collection and multiplication of virus isolates

Leaf samples exhibiting symptoms were collected in four consecutive years (2006–2009) from two geographically distinct areas: from maize (*Zea mays* L. convar. *dentiformis*, *Zea mays* L. convar. *saccharata*), johnson grass (*Sorghum halepense* (L.) Pers.) and grain sorghum (*Sorghum bicolor* (L.) Moench) growing in the experimental field of the Cereal Research Non-Profit Co. in Szeged, and from maize (*Zea mays* L. convar. *dentiformis*, *Zea mays* L. convar. *saccharata*) and johnson grass (*Sorghum halepense* (L.) Pers.) growing in Martonvásár. The samples were stored in RNAlater solution (Ambion) at -70°C until required.

After homogenisation in 0.02 M sodium phosphate buffer (pH 8), the samples were mixed with celite. Plants of the 'Honey' variety of sweetcorn (*Zea mays* L. convar. *saccharata*) in the 2-3-leaf stage were inoculated by rubbing the liquid obtained onto the leaves.

2.2 Pollen viability analysis

The pollen viability of inoculated and control maize plants grown in a phytotron chamber was analysed using FDA staining. Fluorescein diacetate is able to pass through the cell membrane, whereupon intracellular esterases cleave off the diacetate group. The fluorescein accumulates in cells with intact membranes, and can thus be used as an indicator of cell viability. Anthers were removed from the tassels and pollen was collected. The microspores were analysed in 0.3 M mannitol solution containing FDA at a concentration of 0.04 μg ml^{-1} (Widholm 1972).

2.3 Molecular analysis

2.3.1 Sample processing, RNA extraction, reverse transcription

RNA was extracted from the leaves using a Qiagen RNeasy Plant Mini Kit, with an initial quantity of 100 mg. The quality of the RNA extracted was checked by means of gel electrophoresis, and the quantity was determined spectrophotometrically. Reverse transcription was performed using a RevertAid First Strand cDNA Synthesis Kit from Fermentas (M-MuLV-reverse transcriptase) with the oligo dT_{18} nucleotide primer. The reaction mixture was prepared according to the manufacturer's instructions.

Due to the length of the sequences in case of the full-length clone, after treating the cDNA with Ribonuclease H (Ambion) the reverse transcription was performed with the Superscript III RT enzyme (Invitrogen), according to the instructions of the manufacturer.

2.3.2 Polymerase chain reactions

The PCR primer pair (MDMV 8198 fwd 5' AAA CCG GTG GYT RCT YGA ART GD 3'; MDMV3'–5' ATC CTA GGT TTT TTT TTT TTT TTT TTT TTT TTT GTC 3') amplified the 3' non-coding region, the coat protein coding region and an approx. 323 bp segment of NIb (~1317 bp) from the single-strand cDNA.

In all cases an enzyme mixture with a 40:1 Taq:Pfu polymerase ratio was used for the polymerase chain reaction. The cloned virus fragments were sequenced by the Biomi Company, using M13 forward, M13 reverse and amplification primers.

Due to the length of the sequences, the Expand Long Template enzyme mixture (Roche) was used for the full-length clone. The complete virus genome was amplified in two sections, the first from 5' to 4446 (4446 bp) and the second from 4105 to the polyadenylation signal (5400 bp).

2.3.3 Steps in the cloning of coat protein genes

The DNA fragments amplified by PCR and extracted from the gel were inserted into the *AgeI-XmaJI* cloning site of the pLitmus28i (2823 bp, New England BioLabs) vector, which was treated in all cases with phosphatase (Fermentas-SAP-Shrimp Alkaline Phosphatase). Cloning was performed using restriction endonucleases obtained from Fermentas. The ligates were transformed into *Escherichia coli* TOP10 (Invitrogen) competent cells, after which clones containing the insert were identified using blue-white selection. The base sequences of true positive samples were determined using the M13for and M13rev primers.

2.3.4 Transformation methods

The transformation of *E. coli* cells was performed using the method of Inoue (1990). The competent cell suspension was frozen in 50 µl aliquots in liquid nitrogen and then stored at –70°C. After transformation the cells were shaken at 200 rpm in liquid SOC nutrient medium for an hour at 37°C, after which they were spread on solid LB medium supplemented with the antibiotic corresponding to the selection marker. The molecular mechanism for blue-white screening is based on a genetic engineering of the *lac* operon in the *Escherichia coli* laboratory strain serving as a host

cell combined with a subunit complementation achieved with the cloning vector. If the vector construct allowed blue-white selection to be performed, 10 µl 1 M isopropyl-thio-β-galactoside (IPTG) and 40 µl 20 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) was spread on the solid medium, followed by incubation at 37°C overnight. In the case of transformation using the TOPO-XL or TOPO-TA vectors there was no need for blue-white selection, because these vectors contain the lethal ccdB gene at the C terminal end of the LacZ fragment. If the desired sequence is incorporated, this disrupts the ccdB gene, so only bacteria containing successfully ligated vectors are able to survive.

The pAHC20-UbiCPiPC plant transformation vector was introduced into *Agrobacterium tumefaciens* (Ag11) cells using a modification of the Höfgen and Willmitzer (1988) method, whereby it was only grown for one day on solid medium, and the antibiotic concentration was increased to several times the original, due to the low antibiotic sensitivity of the bacterium.

2.4 Vectors, bacteria and media used for cloning

In the course of phylogenetic analysis, the coat protein gene of MDMV was cloned into the pLitmus28i vector (New England BioLabs).

The vectors pLitmus28i (New England BioLabs), pBluescriptKSII+ (Fermentas), pAct1F (McElroy et al. 1991) and pAHC20 (Christensen et al. 1996) were used to prepare a construct containing a truncated MDMV coat protein gene.

The PUC19 (Fermentas), TopoXL (Invitrogen) and TopoTA (Invitrogen) plasmids were used to develop the full-length MDMV clone.

The NOS terminator (nopaline synthase terminator) and the CaMV (Cauliflower mosaic virus) 35S promoter required for cloning were amplified by PCR from the pGWB402 vector (Nakagawa et al. 2007).

2.4.1 Bacterium strains and nutrient media used for cloning

The Top10 (Invitrogen) and JM109 chemically competent cells of *E. coli* and the Ag11 strain of *A. tumefaciens* were used for cloning. The LB (Lauria-Bertani) and SOC media were used for the *E. coli* transformations and the YEB (yeast extract broth, Sambrook et al. 1989) medium for work involving *A. tumefaciens*.

2.4.2 Preparation of competent cells

Chemically competent cells of *E. coli* were prepared using the method of Inoue et al. (1990) and *A. tumefaciens* cells using the method of Höfgen and Willmitzer (1988).

2.5 Bioinformatic analysis

2.5.1 Nucleotide sequence comparison and phylogenetic analysis

Data on the primary structure of the CP genes were evaluated using the BioEdit Sequence Alignment Editor program and then used for the following analyses. The nucleic acid and *in silico* concluded CP amino acid sequence data were used to construct a phylogenetic tree with the ClustalX2.0.9 program. The same results were obtained using three different computing methods (neighbour-joining: NJ; maximum parsimony: MP; maximum likelihood: ML) so the phylogenetic tree was constructed using the NJ results. Bootstrap analysis using 1000 repetitions was applied by the ClustalX program to test reliability. The phylogenetic tree was visualised using the NJplot program (Perrière and Gouy 1996).

The mean diversity and genetic distance (p-distance) were calculated with the MEGA 4.0.2 program (Kumar et al. 2008).

2.5.2 Recombination analysis

Recombination analysis on the newly isolated MDMV sequences and on those available in international data banks was performed using the PDM (probabilistic divergence measures) analysis in the TOPALi v2 program package (Milne et al. 2004) and the RDP3 v.3.44 recombination detection program (Martin et al. 2005, Heath et al. 2006) at the 95% level of significance (window size = 500 nt, step size = 50 nt).

2.5.3 RNA secondary structure analysis

The analysis of RNA secondary structure and RNA stability was carried out using the ViennaRNA (Hofacker et al. 1994), NUPACK (Dirks and Pierce 2004) and UNAFold (Markham and Zuker 2008) programs.

3. RESULTS

3.1 Collection of Maize dwarf mosaic virus isolates and analysis of the symptoms

Samples were collected in the field in Martonvásár and Szeged from maize (*Zea mays* L. convar. *dentiformis*, *Zea mays* L. convar. *saccharata*), johnson grass (*Sorghum halepense* (L.) Pers.) and grain sorghum (*Sorghum bicolor* (L.) Moench).

Despite the fact that the host plants from which the samples were collected exhibited a variety of symptoms, all the sweetcorn (*Zea mays* L. convar. *saccharata*, cv. 'Honey') plants inoculated with the sap of infected leaves exhibited similar symptoms, suggesting that in the case of MDMV the form of symptoms depends on the host plant and environmental factors rather than on minor differences in viral coat protein genes or geographical location.

The presence of MDMV was detected by PCR in a total of 86 cases from samples collected over four years, but no virus was detected using primers specific for SCMV or SrMV, confirming the results of Achon et al. (2007) and Oertel et al. (1997), who found that SCMV was the dominant virus on monocotyledons in Germany, while MDMV was dominant in other European countries.

3.2 Pollen viability analysis

The pollen viability of MDMV-infected and control maize plants grown in phytotron chambers was investigated by means of FDA staining. The viability could not be evaluated as a percentage, as only the cell-walls could be detected in microspores of infected plants; none of the pollen grains were viable.

Due to the dwarfing of the internodes, virus-infected maize plants were only half as tall as the control plants, while the tassels developed two weeks later and were only a quarter of the control size. Microspores could only be obtained from infected plants by dissecting the tassels. No reports were found in the literature on the effect of infection with MDMV, or any other potyvirus attacking monocotyledons, on the pollen production or pollen viability of the host plant.

3.3 Genetic background of Maize dwarf mosaic virus

The nucleic acid and *in silico* concluded CP amino acid sequence data were used to construct a phylogenetic tree describing genetic relationships between the sequences of a total of 94 MDMV isolates (Gell et al. 2010), eight of which originated from the NCBI gene bank database. The latter

had the following registration numbers: MDMV-Hungary ScH/SYN (AJ542536), MDMV-Spain-M (AM110758), MDMV-Spain-Sp (AJ416645), MDMV-A (U07216), MDMV-Argentina-Arg (DQ973169), MDMV-Bulgaria-Bul (NC003377), MDMV-Royalty (AJ563726), MDMV-Bulgaria-Burgas (AM490848).

The mean genetic distance between the isolates was 7.4% (p-distance: 0.074). When various regions of the coat protein gene were examined, it was found that the N-terminal region (1-204 nt) was the most variable, with an average 12.2% deviation in the nucleotide order. The mean sequence deviation of the C-terminal region (717-915 nt) was 7.2% and that of the conserved central region (205-716 nt) 5.2%. The genetic distance of the 3' UTR (252 nt) ranged from 0-11.4% (3.7% on average).

Based on the serological and molecular data of various potyviruses, two main factors have been identified as having a great influence on the coat protein: geographical origin and the host plant (Alegria et al. 2003; Xu et al. 2008). The genetic relationships and genetic distances revealed for the maize dwarf mosaic virus, however, suggest that, unlike the other members of the SCMV subgroup, sequence variation cannot be detected between MDMV isolates on the basis of either geographical location or host plant, and that the virus populations are stable under Hungarian conditions.

The only exceptions are a small number of isolates from Martonvásár (Mv0702, Mv0801, Mv 0811, Mv0814, Mv0905), which carry a 13-amino acid long insertion in the amino-terminal region of the coat protein. One of the Martonvásár isolates (Mv0811, FM883174), collected from johnson grass in 2008, has a 9-amino acid deletion before the insertion. An insertion in the amino-terminal region of the MDMV coat protein was previously described by Tóbiás and Palkovics (2004), but this is the first report of the insertion being accompanied by a deletion. Insertions and deletions are also found in this variable region of the coat protein in other potyviruses, e.g. TuMV/deletion (Lehmann et al. 1997), SCMV/insertion (Oertel et al. 1997), DMV/insertion (Pappu et al. 1994), SCMV (SCE subgroup)/deletion (Alegria et al. 2003), SrMV/insertion (Mirkov et al. 1997). The presence of insertions and deletions in the amino-terminal region of the coat protein confirms the flexibility and adaptability of the genome of potyviruses, and the fact that they occur so frequently suggests that they may possess functional traits. For instance, the virus may be able to bind more strongly to the vector, or spread more rapidly from cell to cell, but no information is yet available to confirm this possibility. Comparative computer based analysis revealed that in some isolates the insertion exhibited a high degree of conservation in terms of both nucleic acid base sequence and localisation. Further computer analysis indicated that the insertion plays a role in virus RNA stability (Petrik et al. 2010).

The isolates containing the insertion formed a separate group on the phylogenetic tree, with the isolate from Argentina. A further point of interest is that in all the isolates carrying the insertion the DAG amino acid motif, representing the aphid transmission factor, changed to DVG. Based on the available sequences, there appears to be a correlation between the insertion and the DAG mutation, though the latter also occurs in isolates without the insertion (MDMV-AJ542536-HUN-Sc-Sc/h).

Several authors have shown that mutations in the DAG motif of potyviruses influences their transmission by aphids (Shukla et al. 1991, Ullah et al 2003, Farreyrol et al. 2006). Atreya et al. (1991) investigated DAG motif mutations in TVMV (Tobacco vein mottling virus) isolates. The question arises of whether the modification of Asp-Ala-Gly to Asp-Val-Gly has any effect on the aphid transmission of MDMV.

In the case of RNA viruses it is a well-known fact that the virus populations may be variable within a given plant due to the constant occurrence of replication errors leading to changes, mutations and recombinations (quasi species). In order to analyse the intraspecific structure of the populations of the virus isolates described in the present work, the sequences of 8 clones were determined for each of two isolates (Mv0702, Sz0612). A sequence homology of 99-100% was found for the Mv0702 isolate, while sequence deviations of 0.1-7.6% were recorded for the Sz0612 sample. At amino acid level, however, this difference was much smaller (0-3.1%). It can thus be said that the populations of the tested isolates were homogeneous within the plant.

3.4 Virus RNA secondary structure analysis

The insertion in the amino terminal region of the coat protein and the mutation in the DAG amino acid motif were also observed in other potyviruses (ZeMV Seifers et al. 2000, MDMV Tóbiás and Palkovics 2004). It appeared that the insertion was the result of duplication in these isolates too, but little is known of its function. Functional analysis has been performed on the RNA 5’ and 3’ secondary structure of a number of viruses (McCormack et al. 2008), and significant deviations were detected between different virus genera on the basis of genome-scale ordered RNA structures (GORS) (Simmonds et al. 2004). These authors concluded that these RNA structures are missing from the genome of potyviruses; the need for these structures is probably limited by interactions between RNA and nucleocapsid proteins. The present work investigated the effect of the insert in the N terminal on the stability and secondary structure of MDMV RNA. The secondary RNA structures of the newly collected MDMV coat protein genes and those stored in the database were estimated using three different programs, the Vienna RNA package (Hofacker et al. 1994), UNAFold (Markham et al. 2008) and NUPACK (Dirks and Pierce 2004). Both the Vienna and the

UNAFold programs employ the Zuker algorithm (Zuker and Stiegler 1981), while NUPACK also allows pseudoknot RNA structures to be analysed.

It can be concluded from the results that isolates containing the insertion had lower minimum free energy (MFE) than those without these sequences. The values after the random sequences (867, 876, 915) designate the length of the coat protein genes, which was 876 bp for Hungarian isolates, 867 bp for Spanish isolates and 915 bp for those containing the insertion. On average, isolates containing the insertion had MFE values 7.7% lower than sequences containing the 915 bp random insertion, and 7.2% lower than that of sequences without the insertion. The ANOVA test was used to calculate significant differences.

The role of RNA secondary structure in virus replication is well known (Simmonds et al., 2004), so it could be that the insertion influences virus replication by stabilising the secondary structure.

Based on the RNA secondary structure of the virus, it can be said that the insertion detected in the N-terminal region of the MDMV coat protein gene stabilises the molecular structure with lower minimum free energy value and less branches of the RNA molecule.

3.5 Recombination analysis

The TOPALi V2 program was unable to detect recombination sites in the maize dwarf mosaic virus coat protein gene when comparing different isolates of the virus, and related viruses. The analysis of the two full-length virus isolates determined in the present work (MV0801-M, Sz-0605-M) and of three full-length MDMV sequences from the database (NC_003377.1, AM110758, AJ001691) revealed a recombination breakpoint at 2951 nt in the P3 protein-coding region. Based on the first segment (1–2951 nt) the isolates can be grouped according to geographical origin, while the second segment (2951–9563 nt) indicates that the Mv0801-M isolate is closely related to the Sp isolate. It is thus worth noting that although the coat protein-coding region is relatively homologous, irrespective of geographical origin and host plant, the recombination point characteristic of MDMV is to be found in the P3 protein-coding region, the most variable region of the potyviruses.

The coat protein-coding regions of the full-length MDMV isolates and the SCMV subgroup were then examined by means of recombination analysis using the recombination detection program RDP3. This program runs nine different algorithms simultaneously (RDP, GENECONV, Boot Scan, MaxChi, Chimaera, SiScan, Phylpro, LARD, 3Seq), each of which uses different algorithm. Nowadays this program is most frequently used in research on recombination variants of HIV (*Human immunodeficiency virus*), but experience shows that it is also ideal for the recombination

analysis of potyviruses. The program identified four previously unknown recombination events in five full-length MDMV sequences (AM110758-Spain, NC_003377-Bulgaria, AJ001691-Bulgaria, Mv0801-M, Sz0605-M). Among the members of the SCMV subgroup, recombination events were found in the coat protein-coding region between the SCMV–SCMV, SrMV–SrMV, JGMV–MDMV (70–210), MDMV–MDMV and SrMV–MDMV (329–397) isolates. Numerous recombination events were detected with the RDP program, while TOPALi identified only a single breakpoint in the full-length MDMV genome.

3.6 Determination of the nucleotide sequence of the whole genome

In order to determine the sequence of the whole MDMV genome, overlapping virus segments were cloned and sequenced. The electropherograms were visualised using the BioEdit Sequence Alignment Editor program and aligned to sequences in the gene back database. The results were used to compile full-length sequences for the isolates Sz0605 (Szeged) and Mv0801 (Martonvásár). The CloneManager program was used to analyse the restriction pattern of the whole genome, which was then used in the preparation of a full-length clone.

3.7 Preparation of a construct capable of inducing resistance to the MDMV virus

The following risk-reducing factors were applied in preparing the construct: the virus sequence is not translated into protein due to the lack of the translation start codon; the N-terminal region of the coat protein does not contain the region coding for the DAG motif; as the latter is essential for virus transmission to aphids, this means that in the case of mixed infection with a virus not transmitted by aphids it will not provide the latter with a new vector; resistance is PTGS-based, so RNA originating from the transgene is not accumulated (Tepfer and Balázs 1997); in the plant transformation vector the ampicillin resistance gene is located outside the boundary sequences, thus allowing positive selection for the bacteria, which are not, however, transmitted to the plants. The construct designed to induce virus resistance works on the principle of gene silencing: the truncated coat protein gene of the virus is present in the hairpin construct in the form of inverted repeat separated by a rice intron. The virus sequences are complementary, so after splicing the intron, double-stranded RNA molecules are formed, from which the MDMV-specific siRNAs act as PTGS elicitors.

Plant transformation and the raising of potentially transformed maize plants is now in progress, but unfortunately maize varieties and lines exhibit enormous differences in regeneration ability, so it will be necessary to elaborate satisfactory transformation and regeneration methods.

3.8 Preparation of a full-length cDNA clone for maize dwarf mosaic virus

A full-length clone of the maize dwarf mosaic virus was prepared from the Mv0801-M isolate. Among the members of the SCMV subgroup, a full-length infectious cDNA clone has only yet been prepared for the *Johnsongrass mosaic virus* (JGMV) (Kim et al. 2003). Infectious cDNA clones have been successfully developed by many authors (PPV Maiss et al. 1992; LMV Yang et al. 1998; TuMV Sánchez et al. 1997; ZYMV Gal-on et al. 1991; PSbMV Johansen, 1996), but there are still many potyviruses where it has not proved possible to induce infection with cDNA clones even after many years of experimentation. The following factors may be responsible for this: the absence or shortness of the polyadenylation signal; errors in one or two bases at the 5' end of the virus, which could be sufficient to cause a complete loss of virulence; the insertions may result in the instability of the virus genome; quality and quantity of DNA; host plant; type of inoculation (manual, *Agrobacterium*-mediated, biolistic).

Unfortunately, even after twelve separate purifications, none of the full-length cDNA clones induced infection on maize plants, all of which remained symptom-free. The temperature and light intensity in the phytotron chamber were optimal for maize growth and the development of viral symptoms, as proved by the positive control plants. When RNA was extracted from plants inoculated with the cDNA clones, no maize dwarf mosaic virus was detected by PCR analysis.

4. DISCUSSION (CONCLUSIONS AND RECOMMENDATIONS)

Maize production is of great significance in Hungary, having an important role in human nutrition and being essential for the feeding of livestock. Due to the rapid spread of bioethanol production, its industrial use is also on the increase. Among the viral diseases attacking maize, MDMV is one of the most important pathogens, as it causes great yield losses. It is thus essential to control weeds that act as a virus reservoir and to protect the crop from aphid vectors.

During the period 2006–2009, PCR analysis of RNA extracted from leaf samples of field-grown maize, sorghum and johnson grass revealed that MDMV is the dominant virus among the members of the SCMV subgroup in the sample locations in Hungary.

The 4-year analysis of the primary structure of coat protein genes of maize dwarf mosaic virus isolates showed that there was no great variability in the genetic material of MDMV. The virus population can be said to be stable, both in Hungary and in other parts of the world.

The studies proved that MDMV infection not only causes reductions in the ear size, grain yield and tassel size, but also has a great influence on pollen viability.

An analysis of the 86 MDMV coat protein genes identified in the present work and of MDMV sequences found in the database demonstrated that there is little chance of achieving a breakthrough in virus resistance on the basis of truncated coat protein genes, since the population proved to be homogeneous and differences occurring in the nucleotide sequence did not in general translate into differences at the amino acid level.

New scientific results

1. In the course of this work the primary structure of the partial NIb and full-length coat protein gene of 86 MDMV isolates was analysed in four consecutive years (2006–2009), thus greatly expanding the collection of MDMV sequences available in the gene bank database.
2. When the isolates were tested on the same host under the same environmental conditions it was found that MDMV isolates which exhibited diverse symptoms under field conditions all produced the same symptoms. The differences could thus be attributed to the influence of the host plant and temperature on the development of virus symptoms and to the phonological stage in which the plant was exposed to infection.
3. The new sequence information and that available in the database was used to prepare a phylogenetic tree for MDMV and its relatives. This was the first analysis of the genetic background of MDMV to be based on the primary structure of coat protein genes originating from such a large number of samples.
4. One of the MDMV isolate (Mv0811, FM883174) coat protein gene containing not only an insertion but also a deletion in the N terminal region was described for the first time.
5. Three different computer programs were used to calculate RNA stability and RNA secondary structure from the primary structure of isolates containing the insertion, and it was concluded that the occurrence of an insertion in the variable N-terminal region of the coat protein gene could influence the stabilisation of the RNA secondary structure.
6. Full-length sequences were determined for two maize dwarf mosaic virus isolates originating from geographically distinct locations (Mv0801 and Sz0605).
7. A construct containing the truncated coat protein gene of the *Maize dwarf mosaic virus* was prepared with the help of gene silencing for use in inducing MDMV resistance.
8. Not only the pollen production but also the pollen viability of sweetcorn plants inoculated with MDMV in the 3-leaf stage of development was found to be substantially impaired. This was the first such observation for the potyvirus infection of monocotyledons.
9. The results of recombination analysis indicated that intermolecular rearrangements are characteristic of the MDMV genome. Recombination events were also found to occur in the coat protein coding region of viruses belonging to the SCMV subgroup.

REFERENCES

ACHON, M.A., SERRANO, L., ALONSO-DUEÑAS, N., PORTA, C. (2007): Complete genome sequences of Maize dwarf mosaic and Sugarcane mosaic virus isolates coinfecting maize in Spain. *Arch. Virol.*,152 (11): 2073-2078.

ALEGRIA, O.M., ROYER, M., BOUSALEM, M., CHATENET, M., PETERSCHMITT, M., GIRARD, J.C., ROTT, P. (2003): Genetic diversity in the coat protein coding region of eighty-six sugarcane mosaic virus isolates from eight countries, particularly from Cameroon and Congo. *Arch. Virol*.,148(2): 357-72.

ATREYA, P.L., ATREYA, C.D., PIRONE, T.P. (1991): Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. *Proc. Natl. Acad. Sci.USA,* 88: 7887-7891.

CHRISTENSEN, A.H., QUAIL, P.H. (1996): Ubiquitin promoter-based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.*, 5:213-218.

DIRKS, R.M., PIERCE, N.A. (2004): An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *J. Comput. Chem*., 25: 1295-1304.

FARREYROL, K., PEARSON, M.N., GRISONI, M., COHEN, D., BECK, D. (2006): Mosaic virus isolates from French Polynesia and the Cook Islands are Dasheen mosaic virus strains that exclusively infect vanilla. *Arch. Virol*.,151 (5): 905-919.

GAL-ON, A., ANTIGNUS, Y., ROSNER, A., RACCAH, B. (1991): Infectious in vitro RNA transcripts derived from cloned cDNA of the cucurbit potyvirus, zucchini yellow mosaicvirus. *J. Gen. Virol*., 72: 2639-2643.

GELL, G., BALÁZS, E., PETRIK, K. (2010): Genetic diversity of *Hungarian Maize dwarf mosaic virus* isolates. *Virus Genes,* 40: 277-281.

HEATH, L., VAN DER WALT, E., VARSANI, A., MARTIN D.P. (2006): Recombination patterns in aphthoviruses mirror those found in other picornaviruses. *J. Virol.*, 80: 11827-11832.

HILL, J.H., MARTINSON, C.A., RUSSEL, W.A. (1974): Seed Transmission of Maize Dwarf Mosaic and Wheat Streak Mosaic Viruses in Maize and Response of Inbred Lines1. *Crop Sci.,* 14: 232-235.

HOFACKER, I.L., FONTANA, W., STADLER, P.F., BONHOFFER, S., TACKER, M., SCHUSTER, P. (1994): Fast folding and comparison of RNA secondary structures (the Vienna RNA Package). *Monaths. Chem.,* 125: 167-188.

HÖFGEN, R., WILLMITZER, L. (1988): Storage of competent cells for Agrobacterium transformation. *Nucl. Acids Res*., 20: 9877.

INOUE, H., NOJIMA, H., OKAYAMA, H. (1990): High efficiency transformation of Escherichia coli with plasmids. *Gene,* 96 (1): 23-28.

JANSON, B.F. AND ELLETT, C.W. (1963): A new corn disease in Ohio. *Plant Dis. Rep.,* 47: 1107-1108.

JOHANSEN, I. E**. (**1996): Intron insertion facilitates amplification of cloned virus cDNA in *Escherichia coli* while biological activity is reestablished after transcription *in vivo*. *Proc*. *Natl. Acad. Sci. USA*, 93**:** 12400-12405.

KIM, K.S., OH, H.Y., SURANTO, S., NURHAYATI, E., GOUGH, K.H., SHUKLA, D.D., PALLAGHY, C.K. (2003): Infectivity of *in vitro* transcripts of Johnsongrass mosaic potyvirus full-length cDNA clones in maize and sorghum. *Arch. Virol.,* 148: 563-574.

KUMAR, S., NEI, M., DUDLEY, J., TAMURA, K. (2008): MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.,* 9 (4): 299-306.

LEHMANN, P., PETRZIK, K., JENNER, C., GREENLAND, A., SPAK, J., KOZUBEK, E., WALSH, J.A. (1997): Nucleotide and amino acid variation in the coat protein coding region of turnip mosaic virus isolates and possible involvement in the interaction with the brassica resistance gene TuRB01. *Physiol. Mol. Plant. Path., (51)*: 195-208.

MAISS, E., TIMPE, U., BRISSKE-RODE, A., LESEMANN, D.-E., CASPER, R. (1992): Infectious in vivo transcripts of a plum pox potyvirus full-length cDNA clone containing the cauliflower mosaic virus 35S RNA promoter. *J. Gen. Virol*., 73: 709-713.

MARKHAM, N.R. AND ZUKER, M. (2008): UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol*., 453: 3-31.

MARTIN, D.P., WILLIAMSON, C., POSADA, D. (2005): RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*, 21: 260-262.

McCORMACK, J.C., YUAN, X., YINGLING, Y.G. ZAMORA, R.E., SHAPIRO, B.A., SIMON, A.E. (2008): Structural Domains within the 3' Untranslated Region of Turnip Crinkle Virus. *J. Virol.*, 82: 8706-8720.

McELROY, D., BLOWERS, A.D., JENES, B., WU, R. (1991): Construction of expression vectors based on the rice actin 1 (*Act1*) 5' region for use in monocot transformation. *Mol. Gen. Genet.,* 231: 150-160.

MIKEL, M.A**.,** D'ARCY, C.J., FORD, R.E. (1984): Seed transmission of maize dwarf mosaic virus in sweet corn. *J. Phytopathol.*, 110: 185-191.

MILNE, I., WRIGHT, F., ROWE, G., MARSHALL, D.F., HUSHMEIER, D., MCGUIRE, G. (2004): TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics*, 20: 1806-1807.

MIRKOV, T.E. AND YANG, Z.N. (1997): Sequence and relationships of Sugarcane Mosaic and Sorghum Mosaic Virus strains, and development of RT-PCR based RFLPs for strain discrimination. *Phytopathology, 87*: 932-939.

NAKAGAWA, T., SUZUKI, T., MURATA, S., NAKAMURA, S., HINO, T., MAEO, K.,TABATA, R., KAWAI, T., TANAKA, K., NIWA, Y., WATANABE,Y., NAKAMURA, K., KIMURA, T., ISHIGURO, S. (2007): Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Biosci. Biotechnol. Biochem*. 71 (8): 2095-2100.

OERTEL, U., SCHUBERT, J., FUCHS, E. (1997): Sequence comparison of the 3'-terminal parts of the RNA of four German isolates of sugarcane mosaic potyvirus (SCMV). *Arch. Virol.* 142: 675-687.

PAPPU, S.S., PAPPU, H.R., RYBICKI, E.P., NIBLETT, C.L. (1994): Unusual amino-terminal sequence repeat characterizes the capsid protein of dasheen mosaic potyvirus. *J. Gen. Virol., 75*: 239-242.

PERRIÈRE, G. AND GOUY, M. (1996): WWW-Query: An on-line retrieval system for biological sequence banks. *Biochimie*, 78: 364-369.

PETI, J. **(**1983): A kukorica csíkos mozaik vírus kártételének vizsgálata 24 kukoricahibriden. *Növényvédelem,* 19: 18-25.

PETRIK, K., SEBESTYÉN, E., GELL, G., BALÁZS, E. (2010): Natural insertions within the N-terminal region of the coat protein of *Maize dwarf mosaic potyvirus* (MDMV) have an effect on the RNA stability. *Virus Genes,* 40: 135-139.

REVERS, F., LE GALL, O., CANDRESSE, T., MAULE, A.J. (1999): New Advances in Understanding the Molecular Biology of Plant/Potyvirus Interaction. *Mol. Plant-Microbe Interact.*, 12 (5): 367-376.

SAMBROOK, J., FRITSCH, E.F., MANIATIS, T. (1989): Molecular cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press.

SÁNCHEZ, F., MARTÍNEZ-HERERRA, D., AGUILAR, I., PONZ, F. (1998): Infectivity of turnip mosaic potyvirus cDNA clones and transcripts on the systemic host *Arabidopsis thaliana* and local lesion hosts. *Virus Res.*, 55: 207-219.

SEIFERS, D.L., SALOMON, R., MARIE-JEANNE, V., ALLIOT, B., SIGNORET, P., HABER, S., LOBODA, A., ENS, W., SHE, Y.M., STANDING, K.G. (2000): Characterization of a novel potyvirus isolated from maize in Israel. *Phytopathology*, 90: 505-513.

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.

SIMMONDS, P., TUPLIN, A., EVANS, D.J. (2004): Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implications for virus evolution and host persistence. *RNA,* 10: 1337-1351.

SUM, I., SEBESTYÉN, E., PAPP, I., LISZT, A. (1979): A kukorica törpe mozaik vírus hatása 15 kukoricahibridre. *Növénytermelés,* 28: 309-315.

SZIRMAI, J. AND PAIZS, L.-NÉ (1963): A kukorica csíkos mozaik betegsége. *Növénytermelés,* 12: 43-50.

SZIRMAI, J.(1968): The occurrence of stripe mosaic disease of maize in Hungary and possibilities of breeding for virus resistance. *Acta Phytopathol. Entomol. Hung.,* 3: 189-198.

TEPFER, M. AND BALAZS, E. (editors) (1997): *Virus-Resistant Transgenic Plants: Potential Ecological Impact*. Versailles & Heidelberg: INRA &Springer-Verlag.

TOLER, R.W. (1985): Maize Dwarf Mosaic, the Most Important Virus Disease of Sorghum. *Plant Dis.,* 69: 1011-1015.

TÓBIÁS, I., PALKOVICS, L. ÉS PERECZES J. (2003): A kukorica törpe mozaik vírus előfordulása csemegekukoricán. *Növényvédelem,* 39 (6): 247-250.

TÓBIÁS, I. AND PALKOVICS, L. (2004): An unusual feature at the N-terminal end of the coat protein of Maize dwarf mosaic virus isolated in Hungary. *J. Phytopathol.,* 152: 445-447.

TOLDINÉ TÓTH, É. (2008): A vírusfertőzöttség hatása egy kukoricahibrid termésére. *Növényvédelmi Tudományos Napok, Absztrakt*. 18.

ULLAH, Z., CHAI, B., HAMMAR, S., RACCAH, B., GAL-ON, A., GRUMAT, R. (2003): Effect of substitution of the amino termini of the coat proteins of distinct potyvirus species on viral infectivity and host specificity. *Physiol. Mol. Plant Pathol., 63*: 129-139.

WIDHOLM, J. M. (1972): The use of fluoresceine diacetate and phenosaphranin for determining viability of cultured plant cells. *Stain Technol.,* 47: 189-194.

WILLIAMS, L.E. AND ALEXANDER, L.J. (1965): Maize dwarf mosaic virus, a new corn disease. *Phytopathology*, 58: 802-804.

XU, D.L., PARK, J.W., MIRKOV, T.E., ZHOU, G.H. (2008): Viruses causing mosaic disease in sugarcane and their genetic diversity in southern China. *Arch. Virol.*, 153(6): 1031-1039.

YANG, S.J., REVERS, F., SOUCHE, S., LOT, H., LE GALL, O., CANDRESSE, T., DUNEZ, J. (1998): Construction of full-length cDNA clones of lettuce mosaic virus (LMV) and the effects of intron-insertion on their viability in *Escherichia coli* and on their infectivity to plants *Arch. Virol.,* 143: 2443-2451.

ZUKER, M. AND STIEGLER, P. (1981): Optimal computer folding of larger RNA sequences using thermodynamics and auxiliary information. *Nucl. Acids Res.,* 9: 133-148.

Papers related to the dissertation

Papers published in scientific journals

Gell, G., Balázs, E. and Petrik, K. (2010): Genetic diversity of Hungarian *Maize dwarf mosaic virus* isolates. Virus Genes. 40: 277-281. **IF: 1,706**.

Petrik, K., Sebestyén, E., **Gell, G.** and Balázs, E. (2010): Natural insertions within the N-terminal region of the coat protein of *Maize dwarf mosaic potyvirus* (MDMV) have an effect on the RNA stability. Virus Genes. 40: 135-139**. IF: 1,706**.

Gell, Gy., Petrik, K., Sebestyén, E. és Balázs, E. (2010). Kukorica csíkos mozaik vírus (MDMV) populáció jellemzése két kukoricatermesztő területen. Növénytermelés. 59 (4): 19-29.

Gell, G., Petrik, K., Balázs, E. (2011). Unique sequence variant in the CP region of a MDMV isolate. Acta Phytopathologica. 46 (1): 11-15.

Gell Gy., Petrik K. , Balázs E. és Divéki Z. (2008): Kukorica csíkos mozaik vírus (MDMV) populációk molekuláris analízise. Növényvédelem. 44 (11): 567-571.

Scientific presentations and posters:

Gell, G., Petrik, K. and Balázs, E. (2008): Genetic diversity of *Maize dwarf mosaic potyvirus* (MDMV) in Hungary. MMT 2008. évi Nagygyűlése, 2008 október 14-17, Keszthely. (poster)

Petrik, K., **Gell**, G., Divéki, Z. and Balázs, E. (2008): Genetic variability and recombination events of *Maize dwarf mosaic potyvirus* (MDMV). IUMS 2008 Istanbul, XIV. International Congress of Virology, 10-15 August 2008, Istanbul. (poster)

Balázs, E., Petrik, K., **Gell**, G., Divéki, Z. (2008): Recombination studies of *Maize dwarf mosaic potyvirus* (MDMV) as an important factor for risk assessment in maize plants. 10th International Symposium on the Biosafety of Genetically Modified Organisms, 16-21 November 2008, Wellington, New Zealand. (poster)

Gell Gy., Petrik K., Balázs E. és Divéki Z. (2008): Kukorica csíkos mozaik vírus (MDMV) populációk molekuláris analízise. 54. NÖVÉNYVÉDELMI TUDOMÁNYOS NAPOK. (oral presentation)

Papers and presentations not closely related to the dissertation:

Bakó, A., **Gell**, Gy., Balázs, E. (2010) Quantification of transgene expression in maize (*Zea mays* L.) throughout the vegetation period. Plant Breeding. 130: 41-45. **IF: 1,026**.

Gell, G., Bakó, A., Spitkó, T., Pintér, J. and Balázs, E.: Monitoring of transgene expression in maize during the vegetation period. PPBA, Experiences with GMP field trials and combating climate change challenges with green biotechnology. July 4-7, 2010-Kolozsvár, Romania.

Gell, G., Bakó, A., Spitkó, T., Pintér, J. and Balázs, E.: Monitoring of transgene expression in maize during the vegetation period. In: 8th International Symposium in the Series RECENT ADVANCES IN PLANT BIOTECHNOLOGY: NEW DEVELOPMENTS IN GREEN GENE TECHNOLOGY. September 1-4, 2009 - Szeged p. 53. (poster)

Pánczél, S., **Gell**, Gy., Mészáros, A., Balázs, E. 2009. Regenerációs és transzformációs kísérletek különböző olajtök (*Cucurbita pepo* L. var. *styriaca*) genotípusokkal. Lippai János – Ormos Imre –Vas Károly Tudományos Ülésszak, Corvinus Egyetem, Budapest, október 28-30. Abstracts pp. 48-49.

5. Acknowledgements

I would like to express my grateful thanks to Dr Ervin Balázs, Member of the Hungarian Academy of Sciences, Head of the Applied Genomics Department of ARI HAS, for his supervision of my work and for providing me with the opportunity to carry out my research in such a well-equipped laboratory.

Sincere thanks are due to Dr Kathrin Petrik for her constant encouragement and assistance.

I would like to thank Dr Zoltán Divéki for his help and valuable advice.

Thanks are also due to Dr László Hornok, Member of the Hungarian Academy of Sciences, for his constant interest in my progress.

I would like to express my thanks to Endre Sebestyén, PhD student, who always found time to help me with questions concerning bioinformatics.

Thanks are due to Dr Annamária Mészáros for her experiments on maize transformation.

I owe my thanks to Dr Tamás Spitkó and to István Pók, PhD student, and to Dr Éva Tóth from the Cereal Research Non-Profit Co., Szeged for their assistance in sample collection.

I would like to thank Dr Éva Darkó for helping with the pollen viability analysis.

I am grateful to Attila Fábián, PhD student, for his help in preparing the electron microscope images.

Grateful thanks are due to our laboratory assistant, Teréz Babirák, for her precise, conscientious technical assistance.

I would like to thank all the staff of the department, especially Ambrus Bakó, Vilmos Soós and Angéla Juhász, for the cordial atmosphere in which I was able to carry out my research.