# Epigenetic regulation of endogenous plant pararetroviruses

Inauguraldissertation
Zur
Erlangung der Würde eines Doktors der Philosophie
Vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
Der Universität Basel

Von
Faiza Noreen
Aus Pakistan

Universität Basel
Friedrich Miescher-Institute for Biomedical Research
Botanical institute, plant health unit.

Basel, 2005

Genehmight von der Philiosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Thomas Hohn

Prof. Manfred Heinlein

Dr. Katja Richert-Pöggeler

Basel, 20.09.05

Prof. Dr. Hans-Jakob Wirz

(Dekan der philosophisch-
Naturwissenchaftlichen Fakultät)

## Acknowledgements

All praises and thanks for Almighty Allah, who is entire source of knowledge and wisdom endowed to mankind and all respects are for His last Prophet Hazrat Muhammad (peace be upon him), who is forever a torch of guidance and knowledge for humanity as a whole.

At, the outset, I would like to express my appreciation to Professor Thomas Hohn for his advice during my doctoral research endeavor for the past three years. His observations and comments helped me to establish the overall direction of the research and to move forward with investigation in depth. During these years, he was always there to offer quite encouragement by his own extraordinary kindness. I just can not thank him enough for his unconditional support.

I would like to express my sincere thanks to Dr. Katja Richert-Pöggeler who played a major role in making me understand the concept of my research. This thesis grew out of a series of dialogues with her. Her insights into the pararetroviruses enriched my research work. I am also thankful to her as she was always there to greet me in my special moments which helped me to feel at home.

I would like to express my gratitude to Prof. Manfred Heinlein for his constructive suggestions in seminars. Many thanks to Prof. Barbara Hohn and Dr. Mikhail Pooggin for their most helpful advice while writing this thesis. I am very grateful for my years at the Friedrich Miescher institute, where I found excellent working environment. The stay in Botanical institute was beautiful with its lovely garden. For this I am thankful to Prof. Thomas Boller, Prof. Andreas Wimken and Prof. Christian Körner for providing lab space and infrastructure.

I take this opportunity to say thanks to all of my current and former lab fellows Hyun-Sook, Rashid, Jamie, Rajeshwaran, Nachelli, Saule and Mathias. Their presence around has made this time very joyful. I specially would like to thank here Monika, Hannes and Sandra for helping me a lot and for being available whenever I was in need.

Many of the acknowledgements of a thesis are done at the last minute. My acknowledgments are no exceptions. It is during this time that one often forgets to mention some of the biggest inspiration, encouragement and support. I want to thank Philip for his support to complete this thesis and for his warm friendship which enabled me a lot in exploring the Basel. I have learned a lot from him and I am really thankful to know him. Life has blessed me with the opportunity to meet Tanya and Fanna, who together showed me how generosity flows so naturally from a sense of being blessed.

I am forever grateful to my parents, whose foresight and values paved the way for a privileged education, and to my brother and sisters who gently offer counsel and unconditional support at each turn of the road. Especially without the support of my brother I would never had reached this point in my life.

Life is not measured in the number of breaths we take, but in the moments that take our breaths away; this thesis is a small tribute to all of those moments.

## Summary

This thesis focuses on epigenetic processes involved in the regulation of gene expression in endogenous pararetroviruses (EPRVs), exemplified by endogenous *Petunia vein clearing virus* (ePVCV-1) and its episomal form, PVCV. Since ePVCV-1/PVCV was found to have features characteristic of retrotransposon and endogenous retroviruses (Richert-Poggeler and Shepherd, 1997), detailed analysis of these retroelements in different systems gives a deep insight to understand the interconnection of these elements and their regulation by the host cellular machinery as described in chapter one.

Chapter two describes the different silencing states of ePVCV-1 in two distinct *Petunia hybrida* lines, “white 138” (W138) and “rose du ciel” (Rdc). Despite of ePVCV-1 integration into the pericentromeric regions of the *Petunia hybrida* chromatin, we found that this position still allows for a low level of transcription that increases with increasing plant age and is higher in W138 than Rdc. To correlate these findings with epigenetic marks, we compared these cultivars in respect to DNA- and histone-methylation and siRNA production. Using bisulfite treatment, ePVCV-1 sequences were found to be methylated at cytosines in all contexts. Astonishingly, however, in both hosts the methylation rate in the non-coding region containing the promoter is relatively low. This might indicate a special ability of the viral promoter to escape complete inactivation by methylation. In Rdc, nearly all histones covering the ePVCV-1 coding region were methylated at lysine 9 of histone 3 (H3K9), a flag for heterochromatin, while in W138 about half of them were of the H3K9- and half of the H3K4-type, the latter representing active chromatin. Interestingly and in accordance with the DNA methylation data, the H3K4/H3K9 ratio was relatively high for the promoter region of both cultivars. The higher H3K4/H3K9 ratio in W138 correlates with an increased rate of ePVCV-1 induction. Furthermore, we show the production of siRNAs of three different size classes (24, 22 and 21 nt) in both cultivars, all of

which are weaker in W138 than in Rdc. Together our observations indicate that W138 is less efficient in silencing of the endogenous viral sequences than Rdc.

In chapter three, I investigated the promoter region of PVCV and determined its ability to direct transcription in transgenic plants. Furthermore, I analyzed the regulatory elements of this particular promoter in comparison with those of other plant pararetrovirus promoters. In particular I studied the functionality of an as-1 like element and its contribution to PVCV promoter expression. Although originally of medium strength, the promoter could be improved to about 50% strength of that of the CaMV 35S promoter by “repairing“ a pair of degenerated as-1 enhancer elements. We show, that the promoter includes upstream and downstream enhancer elements, and that it can be improved considerably by restoring two degenerated as-1 elements.

The concept of creating virus-resistant plants by transformation with genes derived from the pathogen genome is a well-exploited and highly effective procedure to fight viruses as causal agents of diseases in plants (Fichen and Beachy, 1993). Recently it has been demonstrated that RNA interference (RNAi) can be successfully triggered against plant viruses by transient expression of an inverted repeat of target sequences (Pooggin et al., 2003; Tenllado et al., 2004). In chapter four, we use this technique to develop RNA-mediated banana streak virus resistance via TGS and/or PTGS and the method should prevent the outbreak of virus infection upon rare spontaneous induction of endogenous BSV in tissue culture.

Chapter five is a publication in EMBO journal to which I contributed in major ways. This paper describes the production of cloned PVCV originating directly from Petunia plants and from a Petunia gene library. Our findings allowed comparative and direct analysis of horizontally and vertically transmitted virus forms and demonstrated their infectivity using biolistic transformation of a provirus-free petunia species. Some integrants within the genome of *P.hybrida* were found to be arranged in tandem, allowing direct release of virus by transcription. In addition to known inducers of endogenous pararetroviruses, such as genome hybridization, tissue culture and abiotic stresses, we observed

activation of PVCV after wounding. Our data also support the hypothesis that the host plant uses DNA methylation to control the endogenous pararetrovirus.

In a preamble I point out, which part of this paper is based on my own experimentation and interpretation. on to control the endogenous pararetrovirus.

In a preamble I point out, which part of this paper is based on my own experimentation and interpretation.

# Chapter 1: General Introduction

## 1: Retroelements in the eukaryotic genome

The term retroelement can be applied to any genetic element that employs reverse transcription in its replication (Rothnie et al., 1994). Since the discovery of the reverse transcriptase (RNA dependent DNA polymerase; RT) more than 30 years ago (Temin, 1970), it has become clear that retroelements exist in many different forms including viruses and transposable elements (Hull, 1999) and occur in a wide range of living organisms prokaryotes and eukaryotes. Based on the facts that RT of retroelements has common domains with the RNA-directed RNA polymerase of RNA viruses (Xiong and Eickbush, 1990), it has been suggested that RT was an early and critical enzyme in the origin of DNA-based organisms (Heslop-Harrison, 2000). The ability to be externally transmitted to other cells or hosts distinguishes the viral from the nonviral retroelements. According to the recent classification (Fig.1), the class Retroelementopsida (previously called retroelements), include the group of DNA and RNA reverse transcribing viruses represented by the order of Retrovirales and non-viral elements represented by the order of Retrales (Hansen and Heslop-Harrison, 2004). Retrovirales have further three suborders. The Orthoretrovirineae (retroviruses) have RNA genomes and are transcribed into DNA using RT. This proviral (DNA) form integrates into the host chromosome (Brown, 1990) by a virus encoded integrase. This integration is obligatory for viral replication. The suborder Pararetrovirineae (pararetroviruses) is found in both animals (family; Hepadnaviridae: (Ganem and Varmus, 1987) and plants (family; Caulimoviridae: (Pfeiffer and Hohn, 1983). These viruses encapsulate a double-stranded (ds) DNA circular genome and replicate through an RNA intermediate. No integrase function is detected in their genome and

integration is not an obligatory part of their replication, infection and transmission cycle (Hull and Covey, 1995). The third suborder Retrotransposineae (previously called LTR-retrotransposons; (Boeke et al., 1988), includes the families of Pseudoviridae (formerly known as Ty1-copia group) and Metaviridae (formerly known as Ty3-gypsy group). Like Orthoretrovirineae, the integration of Retrotransposineae is an obligatory step in the replication cycle. They can form virus-like particles and for some representatives infectivity has been shown due to the presence of an env gene enabling them to move from cell to cell (Kim et al., 1994; Song et al., 1994).

Non-viral retroelements are called Retrales and comprise the two suborders of Retroposineae (previously called non-LTR retrotransposons; (Xiong and Eickbush, 1990) and of Retronineae. These retrotransposons have less similarity with Retroposineae compared to Retrotransposineae suborder.

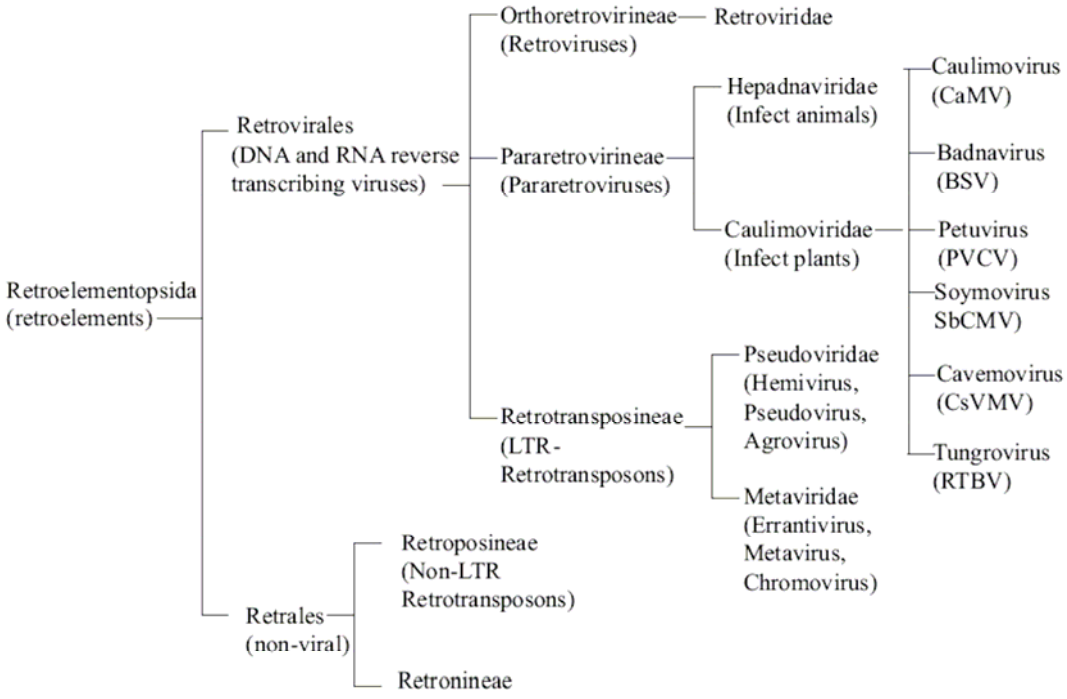


**Fig 1:** Based on recent nomenclature classification scheme of retroelements.

## 1.1: Plant DNA viruses as retroelements

### 1.1.1: Pararetroviruses and their replication

The family of Caulimoviridae with its suborder Pararetrovirineae, contains all plant viruses that replicate by reverse transcription (Fig 2). These viruses have discontinuities at specific sites in their circular double-stranded DNA genomes, one in the antisense and between one and three in the sense strand. These discontinuities are at the priming sites for DNA synthesis and result from initiation of antisense and sense DNA replication. A feature common to the genomic structure of all plant pararetroviruses is the arrangement of the genes encoding structural proteins (gag) and enzymatic functions (pol). The genes are always adjacent like gag-pol. The pol gene shows consensus domains for a protease (PR), a reverse transcriptase (RT) and a ribonuclease H (RNase H) (Rothnie et al., 1994). None of the pararetroviral pol genes encodes an integrase function, as it is typical for pol genes of retroviruses and LTR retrotransposons (like Metaviridae and Pseudoviridae).

Pararetrovirus genome replication includes both nuclear and cytoplasmic phases. Virus particles enter into the nucleus by means of nuclear localization signals at their surface (Leclerc et al., 1999) and release the open circular DNA into the nucleus by an unknown mechanism (Fig. 3).

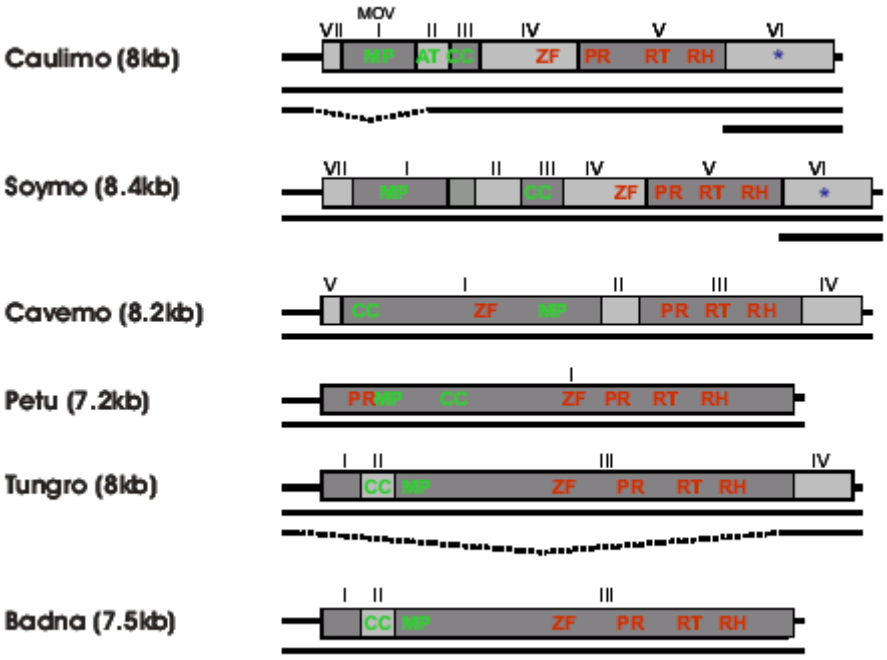


**Fig 2:** Map of the six known genera of plant pararetroviruses, showing the open reading frames (numbered) and indicating the typical motifs (MP, movement protein; CC, coil-coil domain of virion-associated protein; AT, Aphid transmission factor; ZF, Zinc-Finger motif of capsid protein; PR, protease domain; RT, reverse transcriptase domain; RH, RNaseH domain. The RNA produced from the genomes is indicated below the individual maps.

After the virus enters the nucleus, a close-circular episomal DNA is produced by processing the single-strand overhangs, repairing the gaps and ligating the ends using host repair enzymes. For CaMV it has been shown that such DNA molecules together with sequestered histones form minichromosomes (Menissier et al., 1983) that generates full-length viral RNA with terminal repeats (pregenomic RNA, pgRNA) and in some cases also subgenomic RNA. Like in retroviruses, the terminal repeats are produced by concealment of the polyadenylation signal during first encounter with transcription/processing complex and its recognition at second encounter (Sanfacon and Hohn, 1990). However in retroviruses this happens on linear integrants with polyadenylation signals positioned on the two LTRs (Fig. 4), whereas in plant pararetroviruses it occurs with polyadenylation sequences derived from one mono-LTR, encountered twice on circular DNA. It is not clear yet if the transport of viral RNA from the nucleus to the cytoplasm is passive or active. During the cytoplasmic

phase of replication, the interaction of pgRNA with the RNA binding domain of capsid protein subunits results in the formation of virions. Most of virions are embedded in the viral inclusion bodies. This matrix is most likely the place of viral replication in which presumably encapsidated reverse transcribing units are further compartmentalized to be separated from the host cell machinery. Reverse transcription is initiated from a met-tRNA primer and the RNA template is degraded by viral RNaseH activity (reviewed by Hohn and Richert-Poeggeler; submitted). Extension of the minus strand and degradation of the RNA by RNase H produce a terminally redundant minus-strand DNA (Coffin et al., 1997), but polypurine stretches are spared and serve as primers for plus-strand DNA synthesis. Circulization of the genome occurs during plus-strand synthesis and must occur within primer binding site sequence (Bonneville and Hohn, 1993).

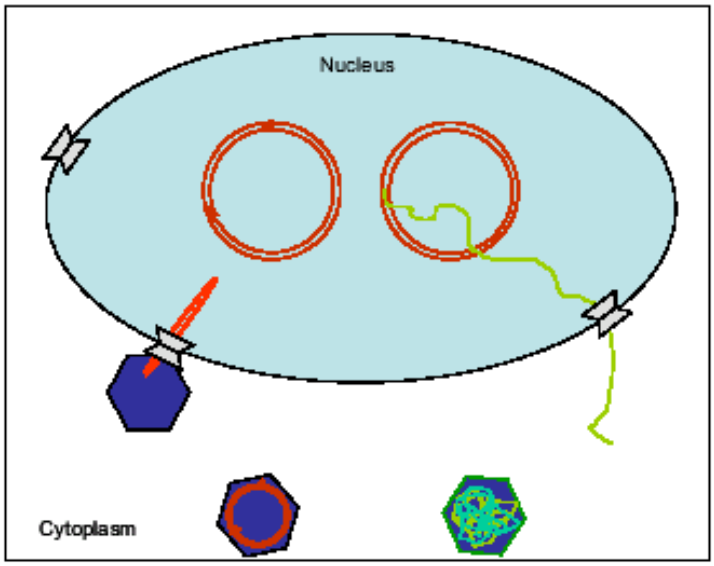


**Fig 3:** Replication cycle of CaMV, typical for plant pararetroviruses.

Pararetroviruses in animals are represented by the hepadnaviruses, known from a small number of mammals and birds. Human hepatitis B virus (HBV) was the first member of the group to be characterized.

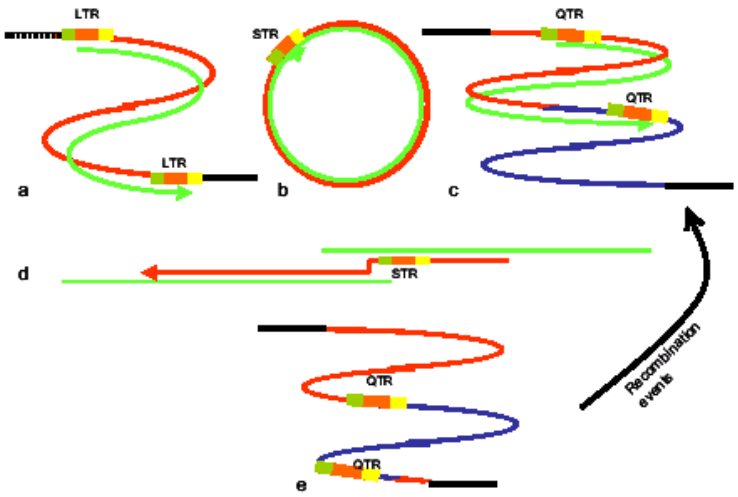


**Fig 4:** Different ways to produce terminally redundant RNA from integrated retrovirus (a), from episomal pararetrovirus genome (b), from head-to-tail dimmer integrants of pararetrovirus genomes by recombination, e.g. during reverse transcription of fragmented RNA (c). (LTR: Long terminal repeat; STR: Solo-long terminal repeat; QTR, Quasi-long terminal repeat).

### 1.1.2: Retrotransposineae (LTR- Retrotransposons)

Recent nomenclature classify also LTR-retrotransposons within the group of Retrovirales (Hansen and Heslop-Harrison, 2004) based on the presence of the env protein that has been identified for several retrotransposons. It is believed now that retroviruses have evolved from LTR-retrotransposons in animals by acquiring the envelope (env) gene (Malik et al., 2000). The structure of LTR-retrotransposons is comparable to that of retroviruses. Commonly, two genes are found in LTR-retrotransposons, representing the homologues of the retroviral gag and pol genes (Malik et al., 2000). Arrangement and functions of these genes largely correspond to those in retroviruses. The env gene of retroviruses distinguishes them from the non viral LTR retrotransposons. Retrotransposineae are further divided into two families of Pseudoviridae and Metaviridae based on sequence similarities of their reverse transcriptase and organization of the subunits within their pol genes.

Pseudoviridae (Ty1-copia group) had initially two genera *Pseudoviruses* (*Saccharomyces cerevisiae Ty-1 virus*; (Hull, 2001) and *Hemiviruses* (*Drosophila*

*melanogaster copia virus*(Hull, 2001). These genera were originally classified on the basis of primer used for reverse transcription: a cleaved half tRNA for *Hemiviruses* or the 3' end of a full tRNA for *Pseudoviruses* (Boeke et al., 2000a, 2000b). However, the recent discovery of Pseudoviridae in plants (*A. thaliana*) with additional env-like genes generated a new genera called *Agroviruses* (Peterson-Burch and Voytas, 2002). The genome arranges as PR, IN, RT and RH in the case of *Pseudoviruses* and *Hemiviruses* while in the case of *Agroviruses* the order of consensus domains is PR, IN, RT, RH and ENV. Pseudoviridae have been extensively studied in the model organisms yeast *Saccharomyces cerevisiae* (*Ty1, Ty2, Ty4* and *Ty5*; (Boeke and Devine, 1998) and *Drosophila melanogaster* (*copia* ; (Potter et al., 1979). Already early data pointed out the insertion sites of Ty1-copia elements close to tRNA genes (Eigel and Feldmann, 1982). Analysis of a large variety of tRNA gene loci showed that the 5' flanking regions of tRNA genes were preferred target sites for transposition (Mewes et al., 1997) indicating a region-specific manner of integration mode. Experimental proof for the tRNA gene regions in yeast (Ji et al., 1993) or the upstream sequences of Pol III transcribed genes in general (Devine and Boeke, 1996) to be preferred target sites for Ty1 transposition has been provided. Transposition rates are low and number of elements is kept constant by balancing transposition and excision events. This may be due to specific interaction of integrase with transcriptional complexes formed over the promoter elements of tRNA (Kirchner et al., 1995).

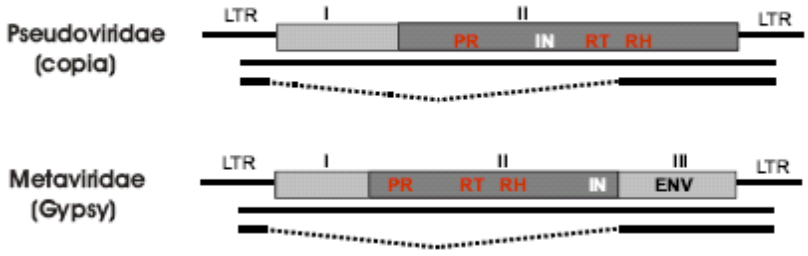


**Fig 5:** Maps of *copia* and *gypsy* transposable elements. The RNA produced from the genomes is indicated below the individual maps.

LTR-retrotransposons like retroviruses undergo intracellular replication by a complex process that first requires synthesis of a pg RNA molecule with terminal repeats (R) that also serves as mRNA. During replication several template switches of RT will guarantee synthesis of complete DNA molecules with long terminal repeats (LTRs) consisting of U3RU5 regions (U3, unique region of the 3'prime end, U5, unique region of the 5'end). The primer binding site (PBS) on the mRNA molecule is complementary to the 3' end of a host tRNA that is used as primer for reverse transcription. RT then synthesizes DNA complement to the R and U5 portions of the 5' end and performs a template switch hybridizing to the homologous sequences of the R region at the 3' end to continue with minus strand DNA synthesis. During synthesis RNase H encoded by the retrotransposon specifically digests the RNA in any RNA: DNA hybrid, thus liberating a single-stranded DNA with homology to the R sequence but leaving RNA sequences necessary for priming. Plus-strand synthesis starts using the polypurine tract (PPT) of the original RNA template as primer towards the 5' end of the already existing DNA strand including the PBS region due to attachment of the tRNA. Finally both RNA primers (t-RNA, PPT region) are digested by RNase H and hybridization between the PBS sequences present in both DNA molecules leads to a temporary circular structure that allows continuation of the reverse transcription until the complete plus strand is synthesized. The circle opens up to allow production of LTRs using the complementary DNA strand as template.

Once the double-stranded linear DNA molecule with LTRs is synthesized, it can be incorporated into the target genome by the action of an integrase, which seems to cut both donor and target molecules. The DNA is apparently cut with nicks that are staggered by 3 to 5 bps (a size that is consistent for any given integrase), thereby creating a flanking target direct repeat that is 3 to 5 bps in size (Boeke and Corces, 1989).

In *A. thaliana* 276 distinct (Boeke et al., 2000b) RTs were identified (Peterson-Burch and Voytas, 2002) which is consistent with the emerging picture of the abundance of transposable elements in plants e.g., *BARE-1* (barley),

*Bs1,Opie-1, PREM-2* (maize),*Tnt1, Tto1* (tobacco) and *Tos17* (rice) (Kumar and Bennetzen, 1999). No *Hemiviruses* were found in plants; however *Agroviruses* including *SIRE-1* of soybean, Endovir of *A. thaliana* and *ToRTL1* of tomato have been identified among *pseudoviruses* (Peterson-Burch and Voytas, 2002). In *A. thaliana Agroviruses* were isolated using specific oligonucleotides derived from the Drosophila *gypsy* env-gene, suggesting that genes evolutionary related to Drosophila *gypsy* Env are also present in retroelements of the plant kingdom (Abdel Ghany and Zaki, 2002). *Pseudoviruses* outside of the plant kingdom are relatively rare (e.g., *S. cervisiae, D. melanogaster*) and for some organisms no Pseudoviridae at all were identified (e.g., nematodes and humans). It is difficult to provide explanation for their distribution. It may be that they originated in plants, where they are ubiquitous, and then moved into other organisms by way of horizontal transfer. For *Agroviruses,* which might have a mean for host cell escape, no horizontal transfer has been reported so far (Peterson-Burch and Voytas, 2002).

The family of Metaviridae encodes the same functions as Pseudoviridae, but with different genomic organization (Fig. 5). The structure of *Metaviridae* resembles that of retroviruses (Xiong and Eickbush, 1990) placing the integrase domain downstream of the reverse transcriptase region. Some Metaviridae like the *gypsy* element of *D. melanogaster* contain an env gene like found in retroviruses that enables viral cell to cell movement. Indeed it had been proven that the *gypsy* element is under certain circumstances infectious and therefore has the ability to function as a retrovirus (Kim et al., 1994; Song et al., 1994). Until recently Metaviridae were split in two genera according to the presence of the env gene (genus *Errantivirus*es with the *gypsy virus* of *D. melanogaster* as type member) or its absence (genus *Metavirus*, type member *Ty-3 virus* of *S. cerevisiae*; (Hull, 2001). *Metaviruses* (*Ty3*) were initially grouped (Malik and Eickbush, 1999) together with the chromodomain-containing elements (*Skipper*; chromodomain at the C-terminal end of integrase). However further studies based on RT analysis places *Skipper* outside the *Metaviruses*. Therefore the

creation of a third genus named *Chromoviruses* e.g., *Gloin* in *A. thaliana* (Marin and Llorens, 2000) has been proposed.

In plants Metaviridae are abundant as well e.g., *Athila*, *Athila 1-1* (*A. thaliana*), *Cereba* (barley), *Cinful-1*, *Grande-1*, *Zeon-1* (maize), *RIRE3* (rice) (Kumar and Bennetzen, 1999).

## 2: Plant retroviruses and their evolution

The possibility that retroviruses might exist in plants had always been addressed (Kumar and Bennetzen, 1999), but with the identification of retrotransposinae containing the *env* gene, evidence has been provided that indeed plant genomes contain retroviral-like sequences (Table 2). The presence of an *env*-like gene that encodes a transmembrane protein is generally considered to be a prerequisite for infectivity of a retroelement (Peterson-Burch et al., 2000). Analogous elements of plants were isolated using specific oligonucleotides for the *gypsy env*-gene, suggesting that *env*-like genes are ubiquitous in the plant kingdom, and are evolutionary related to the Drosophila *gypsy env*-gene (Abdel Ghany and Zaki, 2002). Further studies will be necessary to demonstrate that these putative *env*-like genes encode envelope-like proteins that are capable of transferring retroviral nucleocapsids from cell-to-cell, as shown for the *gypsy* retrotransposon (Kim et al., 1994; Song et al., 1994).

| **Retrovirus-like element** | **Plant** | **Reference** |
|---|---|---|
| SIRE-1 | Soybean | (Laten et al., 1998) |
| Tat1 | A. thaliana | (Wright and Voytas, 1998) |
| Athila4 | *A. thaliana* | (Wright and Voytas, 1998) |
| Cyclops | Pea | (Chavanne et al., 1998) |
| Bagy-2 | Barley | (Vicient et al., 2001) |
| GM-5, GM-6 | Gossypium | (Abdel Ghany and Zaki, 2002) |

**Table 1:** Examples of Plant retrovirus-like elements

What is the origin of plant retroviruses? Phylogenetic analysis of reverse transcriptase sequences of the invertebrate retroviruses strongly suggests that they are derivative of *gypsy*-like retrotransposons (Malik et al., 2000). However, it is still unclear whether acquisition of env genes represents a single event or multiple events (Eickbush, 2002). Interestingly, there are no reports on the presence of *env*-like sequences in of *D. melanogaster* or any other invertebrate or vertebrate (Eickbush, 2002). The presence of *env*-like sequences in both Pseudoviridae and Metavirdae of plants suggests that these two groups of retrotransposons acquired the *env* gene independently. Alternatively, closely related relatives of retroviral derivatives invaded the genome of plants and subsequently lost their env gene (Kumar and Bennetzen, 2000). Currently, it is unknown which process proceeded first. Nevertheless, the existence of plant retroviral sequences supports the hypothesis for an apparent horizontal transfer of viruses in plants (Peterson-Burch et al., 2000).

## 3: Endogenous plant pararetroviruses (EPRV)

Until recently, it was thought that plant viruses, unlike animal and bacterial viruses, did not integrate into the host genome. Observations over the past few years have broken this tent, and now an increasing number of integrated plant DNA viral sequences- the single- stranded DNA geminiviruses and the double-stranded DNA pararetroviruses are being found in plant genomes (Harper et al., 2002). Furthermore it has been found that sequences homologus to those of a non-retro RNA virus, *Potato virus Y* (PVY) are integrated into the genome of several grapevine varieties most likely by recombination between viral RNA and the RNA of host cell retrotransposable element (Vicient et al., 2005). However, this integration into the host genome is passive and not actively promoted by a viral encoded enzyme like an integrase. No obligatory integration of plant viruses has been observed up to now. In all reported cases the integrations are non obligatory, since viral amplification occurs also in host plants that do not contain integrated sequences.

The integrated pararetrovirus sequences have been found for several genera to date, i.e. *petu-, cavemo-, badna- and tungro-viruses* (Jakowitsch et al., 1999; Lockhart et al., 2000; Richert-Poggeler et al., 2003; Gregor et al., 2004; Kunii et al., 2004). The current model of invasion involves viral replication, since the majority of viral junction sequences were found within the gap regions of the viral genome that initiate reverse transcription (Jakowitsch et al., 1999; Kunii et al., 2004). The amplification of "endogenous plant pararetroviruses" (EPRVs) within the plant genome is probably based on several independent integration steps and/or occurred via reverse transcription of integrants (similar to retrotranspososns) followed by recombination between repetitive sequences.

Integrated viral sequences in plants have various degrees of degeneration, depending on the evolutionary time elapsed since the integration event. Naturally, the ones derived from more recent integration events can give rise to episomal forms. A single integrant is not expected to escape by transcription directly, since such an escape would require terminal repeats (LTRs) of the form present in retroviruses. The induction of episomal forms depend on the number of integrants, their completeness, and of course is more easily achieved, if tandem repeated integrates are available.

Three examples of inducible endogenous viruses have been described: *Petunia vein clearing virus* (PVCV; (Richert-Poggeler and Shepherd, 1997) *Banana streak virus* (BSV) (Harper et al., 1999b; Ndowora et al., 1999) and *Tobacco vein clearing virus* (TVCV) (Lockhart et al., 2000).

### 3.1: Banana Streak Virus (BSV)

BSV, a member of the *badnavirus* genus, is the casual agent of viral leaf streak disease of banana (*Musa* spp.) (Lockhart, 1986). The virus is transmitted by the citrus mealybug, *Planococcus citri* (Lockhart and Olszewski, 1993) and possibly some other mealybug species (Lockhart et al., 2000). The wild progenitors of domesticated banana are *Musa acuminata* (A genome) and *Musa balbisiana* (B genome). BSV is very variable, both in its DNA sequence and serological properties (Geering et al., 2005). The genome of only one isolate of BSV from cv. 'Obino l'Ewai (BSOEV; previously known as BSV) has been

completely sequenced (Harper and Hull, 1998), although partial genomic sequences are available for isolates from cvs 'Mysore', 'Goldfinger' and 'Williams' (Geering et al., 2005). BSOEV has severely hindered international banana (*Musa* spp.) breeding programs. New hybrids between A and B genomes are frequently infected with this virus which is activated during tissue culture. Such infection is believed to arise *de novo* from viral sequences integrated in the "B" genome (Harper et al., 1999b; Ndowora et al., 1999) as in the "A" genome Cavendish bananas, the economically most important variety, BSOEV infection is rare, despite the propagation by tissue culture in large quantities for many years (Geering et al., 2001). Viral genomes contain three ORFs in the arrangement typical of *badnaviruses* (Fig 2). As integrated forms viral sequences consisting of two segments, which together comprise the full complement of the virus genome, interrupted by a 6kb 'scrambled region' containing non-contiguous and inverted viral sequences, have been isolated. The current model for activation involves two homologus recombination steps, leading to excision of the 'scrambled region' and the joining of either end of the integrant to give rise to a circular molecule, the equivalent of the virus mini-chromosome (Fig. 4) (Ndowora et al., 1999). Furthermore, integrated BSOEV sequences can be activated in tissue culture (Ndowora et al., 1999), a propagation practice used to multiply planting stocks once a hybrid is made. In order to develop a strategy to fight this harmful disease in banana, we (in collaboration with M.-L. Caruana CRIRAD, Montpellier) tested RNAi against BSV in infected Cavandish banana plantlets.

### 3.2: Tobacco endogenous Pararetrovirus (TEPRV)

The EPRV family in *N. tabacum* was first identified in a routine study to characterize plant-flanking transgenes in *N. tabacum* cv. Petit Havana SR1. One of the isolated sequences showed higher homology to RT sequences of pararetroviruses than to known retrotransposons. This suggested that a viral sequence had been integrated into this region of the *N .tabacum* genome. Additional screening of DNA library provided several sequences with homology to cavemoviruses from which in silico a complete viral genome was assembled (Jakowitsch et al., 1999). Further analysis revealed that the *N. tabacum* genome

harbored up to 1000 copies of the tobacco pararetrovirus-like (TPVL) sequence, which was eventually renamed tobacco endogenous pararetrovirus (TEPRV) (Mette et al., 2002). All sequenced TEPRV copies were defective in one or the other way, suggesting that they are unable to encode functional viral proteins. Integration of TEPRVs into plant chromosomes probably occurred randomly by illegitimate recombination, perhaps at recombinogenic gaps in the open circular form of viral DNA (Jakowitsch et al., 1999). TEPRV sequences are poorly transcribed, probably because they are highly methylated, particularly in CpG and CpNpG nucleotide groups (Mette et al., 2002). Because TEPRVs are present in healthy plants showing no signs of virus infection and the corresponding free virus has yet to be isolated, it has been proposed that they might confer resistance to the exogenous form of virus, perhaps through a gene silencing mechanism involving DNA methylation (Jakowitsch et al., 1999; Mette et al., 2002).

Distinct EPRV families are present in *N. sylvestris* (*Ns*) and in *N. tomentosiformis* (*Nto*), the two diploid progenitors of allotetraploid *N. tabacum*. The NsEPRV family appears identical in *N. sylvestris* and *N. tabacum*, indicating little change has occurred in either species science polyploid formation. By contrast, the NtoEPRV family is larger in *N. tomentosiformis* than in *N. tabacum*, suggesting either preferential elimination from the polyploid genome or specific accumulation in the diploid genome following polyploidization (Mette et al., 2002).

### 3.3: Petunia vein clearing virus (PVCV)

Once induced**,** PVCV produces virions and inclusion bodies similar to those described for cauliflower mosaic virus (Lesemann and Casper, 1973). The PVCV genome has motifs characteristic for Caulimoviridae (Richert-Poggeler and Shepherd, 1997). Phylogeneticallty it is distinct from the five other genera of Caulimoviridae and no serological relationship with other members has been detected. Characteristic for this genus is that the whole coding information is present as one large open reading frame within the viral genome (GenBank Accession No. U95208). Towards the C terminus of the large polyprotein, two amino acid sequence motives (HHCC and DD) (Peterson-Burch et al., 2000)

resemble the catalytic domain of integrases (Richert-Poggeler and Shepherd, 1997), which might reflect degeneration of a once functional enzyme. No further sequence homology to recently identified putative integrase domains of retroelements could be found.

Symptoms vary from mild vein clearing to severe yellowing and epinasty. These symptoms are seen particularly when the plants are under stress, for example water and nutrient deficiency (Lockhart and Lesemann, 1997) and also wounding (Richert-Poggeler et al., 2003).

So far no viral transmission vector could be identified and may only exist in regions petunia plants originated. Therefore, viral infection observed in commercially available hybrid petunia occurs by *de novo* activation from integrated copies. Especially integrants in form of head-to-tail concatamers within host chromosomal DNA with domains similar to the long terminal repeats of retroviral proviruses could mediate direct transcription of full-length copies (Richert-Poggeler et al., 2003). Sequence homologies of identified episomal and chromosomal PVCV sequences were high of about 95%.

Interestingly, the integrated PVCV sequences were found predominantly in the pericentromeric regions of the petunia chromosomes (Richert-Poggeler et al., 2003), which consist of heterochromatin. Clusters of integrated PVCV DNA have been identified by fluorescent *in situ* hybridization (FISH) in five of the seven *P.hybrida* (RDC and W138) chromosomes. The fluorescent signals generated are high in abundance and in strength. This high-copy-number PVCV integration pattern in petunia is distinct from that of the BSV/banana system (Harper et al., 1999b). In petunia, the inactivity of certain transgenes has also been correlated with their integration into heterochromatin (Prols and Meyer, 1992). From the frequency of occurrence in the genomic library, we estimate that about 50–100 copies of integrated viral sequences exist in the haploid petunia genome. Despite this high frequency, spontaneous release of viruses is rare. However, this release can be induced for instance by wound stress (Richert-Poggeler et al., 2003). It has been proposed that wound stress would cause callus formation accompanied by chromotin replication and transient hypomethylation (Richert-

Poggeler et al., 2003). Despite the fact that integrated PVCV sequences of healthy plants are hypermethylated, release of virus is possible. It has been proposed that upon induction, only a short period of demethylation in a single plant cell at one of the ~100–200 integration loci would be enough to release the virus (Richert-Poggeler et al., 2003). For retrotransposons, it has been shown that stress conditions induce transcriptional activation (Grandbastien, 1998b).

## 4: Endogenous retroviruses (ERVs) versus endogenous pararetroviruses (EPRVs)

Retrovirus replication has two unique and characteristic properties; reverse transcription of viral RNA into DNA, and integration of newly synthesized DNA into the genome of the infected cell (Boris-Lawrie and Temin, 1994). Retroviruses usually only infect somatic cells, and consequently retroviral genes integrated into genomic DNA are not passed on to host progeny. Some types of retrovirus, however, can also occasionally infect germ line cells, thereby colonizing the host germ line (Vogt, 1997). Offspring that develop from infected germ line cells will carry the integrated retrovirus as a part of their genomes, and these retroviruses can therefore be subsequently transmitted vertically from one host generation to the next. Retroviruses that enter the germ line in this way are referred to as endogenous retroviruses (ERVs) to distinguish them from horizontally transmitted, exogenous retroviruses (Vogt, 1997). It is possible that many ERVs entering the gene pool are strongly negatively selected and hence do not remain in the host population for long (Stoye, 2001). However ERVs that survive may be transmitted vertically to host progeny. Endogenous retroviruses, such as *gypsy*, are retroviruses integrated in the chromosome but activated by nuclear gene to produce RNA copies and infectious virus (Suoniemi et al., 1998). Human endogenous retrovirus (HERV) is an other example of potentially active ERV(Turner, 2000).

EPRVs are similar to ERVs in replicating by RNA copies that are then reverse transcribed to make DNA. The DNA integrates into chromosome of infected organism, and that can be further activated to release infectious virus. The PVCV was found to have features characteristic of retrotransposon including

DNA sequences for an integration enzyme and other features similar to *gypsy* (endogenous retrovirus) (Richert-Poggeler and Shepherd, 1997).

## 5: Petunia: history and research

The genus *Petunia*, established by Jussieu in 1803, comprises about 30 (sub) species and belongs to the family of *Solanaceae*. Its main geographical distribution is from Argentina to Uruguay and in the Southern part of Brazil as well as in the Andean foothills (Wijsman, 1982; Sink, 1984; Ando, 1996). *Petunia axillaris* (*Petunia nyctaginiflora*) was first cultivated in 1823, and *Petunia integrifolia* (*Salpiglossis integrifolia*) first flowered in the Glasgow Botanical Garden (UK) in July 1831 (Bailey, 1986). Petunia is considered to be the first cultivated bedding plant and has remained one of the favorite genera for developing new varieties. It was not until the 1950s that geneticists began to try to predict new color classes from their genetic and biochemical analyses on *Petunia*; until then research had been fairly frugal and practice-driven.

The first hybrid petunia is thought to have been created about 200 years ago by artificial crossings between the purple flowering *P. integrifolia* and the white flowering *P. axillaries* (Sink, 1984). Since then several breeding steps followed the first artificial crossings between wild type petunias to generate nowadays broad selection of hybrid petunia.

There are three hybrid varieties with common use in the scientific society: the Mitchell variety, which is a doubled haploid from a complex hybrid between *P. axillaris* and the cultivar 'Rose of Heaven' that exhibits superior fertility, growth, tissue culture and transformation abilities; the line V26, a bluish purple line that has been used for antisense and cosuppression studies, flavonoid gene isolation and ethyl methane sulfonate mutagenesis; and the line W138 ,which is practically untransformable but is renowned for its active endogenous dTph1 transposable element system and which has already produced many interesting mutants (Gerats and Vandenbussche, 2005). Rdc and W138 were the two *P. hybrida* cultivars used in this study to analyze the epigenetic control of the endogenous plant pararetrovirus petunia vein clearing virus (ePVCV).

A number of petunias harbor one retroelement, ePVCV, which combines features of both viral and non-viral retroelements (Richert-Poggeler et al., 2003). Amplification of episomal PVCV induces typical vein clearing symptoms. Integration of PVCV sequences has been preserved in the pericentromeric region of petunia chromosomes and probably does not compromise expression of important plant genes because its own expression is repressed by its heterochromatin position (Richert-Poggeler et al., 2003). Whereas all *P. hybrida* analyzed so far contained ePVCV in their genome, some of the wild type petunia were proviral free (*P. parodii*, *P. inflata*).

## 6: Host control over endogenous viral sequence

Like almost every eukaryote, plants have the potential to neutralize invading nucleic acids to prevent deleterious effect on genome stability and expression. Endogenous and exogenous invaders such as transposable elements, viruses and transgenes can be transcriptionally silenced (transcriptional gene silencing, TGS) through histone methylation, DNA methylation or chromatin remodeling (Almeida and Allshire, 2005; Kanno et al., 2005; Kawasaki and Taira, 2005). These invaders can also be post-transcriptionally neutralized (post-transcriptional gene silencing, PTGS) through RNA degradation. Togather these phenomena's are referred as RNA silencing.

### 6. 1: RNA silencing

The term RNA silencing summarizes "RNA mediated gene silencing" in which short RNA molecules trigger repression of homologous sequences. It is a highly conserved pathway, found in large variety of eukaryotic organisms, and its main characteristic is the use of small RNA molecules of 21-26 nucleotides that confer high specificity to the target sequence. Originally, it was described as part of co-suppression' phenomenon in plants (Napoli et al., 1990) or 'quelling' in *Neurospora crassa* (Cogoni et al., 1996) and was later attributed to a posttranscriptional gene silencing process (PTGS) occurring in the presence of complementary RNA molecules that would bind and form double-stranded RNA (Metzlaff et al., 1997). A closely related effect described in *Caenorhabditis*

*elegans* as ‘RNA interference’ (RNAi) (Fire et al., 1998; Ketting and Plasterk, 2000) also requires long double-stranded precursor RNAs to induce and sustain efficient posttranscriptional repression of homologous sequences. The so far described RNA silencing phenomena of PTGS (so-suppression), quelling and RNAi happen in the cytoplasm where, double-stranded RNA (produced by various mechanisms) enters the ‘canonical pathway’ after cleavage into small (21–26 nts) RNA duplexes by an enzyme named the helicase/RNase-like III *Dicer* with functional domains similar to those of helicase and RNase III (Bernstein et al., 2001) . Following ATP dependent unwinding, a single-stranded small RNA (small interfering RNA: siRNA) becomes part of protein complexes in which PAZ/PIWI domain proteins (PPD or *Argonaute*) are central components players (Fagard et al., 2000; Hammond et al., 2000). These RNA-induced silencing complexes (RISC) then target homologous mRNAs and exert silencing either by inducing cleavage (‘slicing’) or, as in the case of micro-RNA-loaded RISC (see below), by also eliciting a block to translation.

RNA-dependent RNA polymerase (RdRP), represented by a multiple gene family in plants (Dalmay et al., 2000; Mourrain et al., 2000)with key functions in various silencing pathways, also plays a role in nematodes (Smardon et al., 2000) and fungi (Cogoni and Macino, 1999; Volpe et al., 2002), but is apparently not required or detectable in the genomes of flies and vertebrates. RdRP amplifies the RNAi/PTGS response by generating more double-stranded RNA from single-stranded targets that can then enter and continue to stimulate the RNA silencing pathway. This positive-feedback system is crucial in plants and worms to amplify the siRNA signal transmitted from cell to cell and to mount a systemic form of silencing (Palauqui et al., 1997; Vaistij et al., 2002).

It is now evident that the core machinery required for RNA silencing plays crucial roles in cellular processes as diverse as regulation of gene expression (cytoplasm), protection against the proliferation of transposable elements and viruses (nucleus and/or cytoplasm) and modifying chromatin structure (nucleus). While it appears that the basic pathway has been conserved, specialization has

adapted the common RNA silencing machinery for these different purposes. This specialization is most obvious in plants, which also encode multiple RdRP and Dicer-like proteins that are relevant for distinct small RNA pathways (Xie et al., 2004) and are acting in different cellular compartments.

### 6.2: Endogenous sequences and heterochromatin

Heterochromatin has been first described 77 years ago y the botanist Emil Heitz (director of the Botanical Institute at the University of Basel from 1937-1955) who observed that a certain proportion of nuclear material stayed condensed during the cell cycle and therefore stained differently in cytogenetic studies.

Transposable elements (TEs) are often found in heterochromatin that is transcriptionally silent and structurally distinct from the open euchromatin (Martienssen and Colot, 2001). These silenced regions have conspicuous features, which can include dense methylation of DNA (5-methylcytosine; 5-Me-C), hypo-acetylation of lysine residues in the N-terminal tails of histones H3 and H4 and methylation of specific lysine residues such as lysine 9 on histone H3 (H3K9me2/3). Some of these modifications create binding sites for particular proteins that, in general, promote transcriptional repression and the formation of silent chromatin or heterochromatin (Fujita et al., 1999; Lachner et al., 2001). The packaging of TEs into heterochromatin represses their expression and blocks their ability to transpose. Furthermore, the assembly of TEs into this ‘silent’ chromatin is an effective way of inhibiting TE proliferation that has been employed by many eukaryotes. Because this form of regulation based on chromatin structure is independent of the primary DNA sequence, specialized mechanisms for recognizing these parasitic elements must be required to selectively trap them in heterochromatin. It is now evident that siRNA are signaling molecules in the sequence specific formation of heterochromatin.

### 6.3: RNA silencing reaches chromatin

The same principle that acts to repress genes posttranscriptionally can enforce modification of homologous chromatin in a way that alters its structure

and consequently its function. RNA-directed DNA methylation (RdDM). One form of RNA silencing in the nucleus (Kanno et al., 2005), was first described in viriod (small, circular single stranded RNA molecules) infected plants (Wassenegger et al., 1994). Furthermore RdDM of promoter sequences referred to as transcriptional gene silencing (TGS) was initially observed in plants and was associated with repression of exogenously introduced transgenes and viral suppression (Mette et al., 2000). Remarkably, the presence of dsRNAs homologous to the promoter or the coding region in the DNA result in silencing that persists even after the trigger has been removed (Mette et al., 2000; Jones et al., 2001). The TGS response triggered by double-stranded RNAs results in the complete transcriptional shutdown of a gene and is associated with *de novo* DNA methylation on the homologous DNA sequences.

TGS indeed appears to be employed to silence/inhibit the activity of several classes of TEs in plant genomes. Apart from essential proteins like Dicer and Argonaute for the processing and targeting of small RNA molecules, the persistence of TE DNA methylation in *Arabidopsis thaliana* requires chromatin-modifying factors such as histone deacetylases, methyltransferases, DNA methyltransferases and *SWI2/SNF2*-related chromatin remodeling components – some of which are also required for the persistence of TE siRNAs (Lippman et al., 2003) and for PTGS (Murfett et al., 2001). This underscores the intimate relationship between RNA silencing and chromatin regulation in plants and their role in repression of TEs and other endogenous sequences (Lippman et al., 2004; Probst et al., 2004).

Small RNAs are also known to direct chromatin modifications in other organisms. For instance, in the ciliate *Tetrahymena thermophila*, small RNAs are used to mark particular DNA sequences for elimination from the transcriptionally active macronucleus, most of which are of a repetitive nature (Mochizuki and Gorovsky, 2004). In the fission yeast *Schizosaccharomyces pombe*, it has been clearly demonstrated that RNA silencing acts to facilitate chromatin modifications over repetitive sequences for the purpose of TE silencing, as in plants, but also

impacts upon basic chromosomal functions (Schramke and Allshire, 2003; Volpe et al., 2003).

### 6.4: Formation of dsRNA is the trigger for RNA silencing

Initial studies in *C. elegans* revealed the importance of dsRNA within the silencing pathways (Fire et al., 1998) and later it was identified as precursor for siRNA production (Sijen et al., 2001).

There are several possible ways by which TEs and repeat dsRNA are generated.

A) Two transcription events on opposite strands converging on any given sequence could generate complementary transcripts that would combine and form dsRNA (Fig. 6A). Alternatively, complementary strands could be transcribed from different copies residing at distinct locations in the genome and subsequent hybridization would allow the formation of a dsRNA substrate.
B) Another simple way of obtaining dsRNA is by transcribing an inverted repeat, which produces a transcript that can form a stem–loop or hairpin structure (as with endogenous microRNA precursors: Fig. 6B).
C) An alternative explanation is that RdRP can in some ways recognize transcripts coming from TEs or viral sources as ‘aberrant’ or ‘foreign’ transcripts and use them as templates to generate dsRNA (Xie et al., 2001). This is supported by the observation that RdRP can produce dsRNA in vitro from ssRNA template in a primer-independent manner (Makeyev and Bamford, 2002). More recent work suggests that transcripts lacking a 5’ cap are targeted by RdRP, although it is still possible that RdRP is attracted to other characteristics, such as premature termination or absence of polyadenylation, or a combination of these features (Matzke et al., 2001)
D) In the case of repeat arrays, such as those commonly associated with pericentromeric regions, it has been suggested that this arrangement results in the production of transcripts that serve as more efficient RdRP

substrates, thus ensuring the stability of the assembled heterochromatin over these regions (Martienssen, 2003).

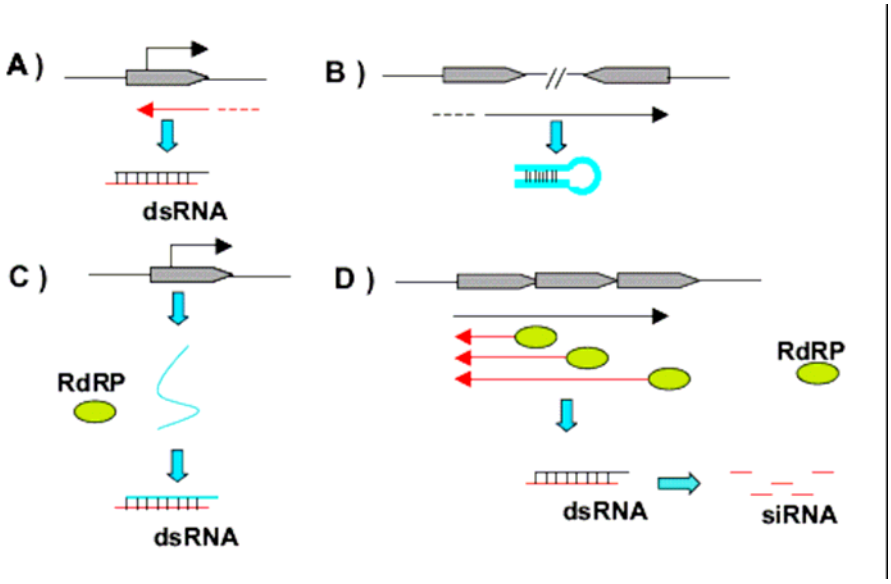


**Fig 6.** Possible routes of dsRNA generation (for details see text).

## 7: DNA methylation and heterochromatin correlation

A common epigenetic modification of DNA in vertebrates, plants and some fungi is the addition of methyl group at the 5$^{th}$ position of cytosine residues. DNA methylation is an enzymatic reaction carried out by proteins of the DNA methyl transferase family. DNA methylation in plants and vertebrates is essential for gene regulation in early embryogenesis, genomic imprinting and protection of genome against migrating transposable elements and retroviruses (Meehan and Stancheva, 2001; Bird, 2002). In a more general sense, DNA methylation can be viewed as an obligatory component of transcriptionally silent chromatin.

Although methylation is the major modification of DNA in the eukaryotic genome, histone molecules can be post translationally modified by methylation, acetylation, phosphorylation, ADP-ribosylation and ubiquitination resulting in histone code. (Peterson and Laniel, 2004). The impact of histone modifications on chromatin structure and dynamics depends on the number of attached modified groups on the position of modified amino acid, and on the histone

protein on which the modified amino acid is located (Jenuwein and Allis, 2001; Lachner and Jenuwein, 2002). In a wide variety of species, acetylation and methylation of selected lysine groups in histone 3 (H3) and histone (H4) amino-terminal tails seem to be crucial for the formation of functionally distinct chromatin states referred to as heterochromatin and euchromatin (Kouzarides, 2002; Richards and Elgin, 2002). These two forms of chromatin differ in their structural and biochemical properties, and carry specific histone and DNA modifications. Thus, for example, trimethylation of lysine 4 (K4me3) and dimethylation of arginine 17 (R17me2) of histone H3, or acetylation of lysine 16 (K16ac) of H4 are modifications that are enriched at transcriptionally active euchromatin loci (Turner, 2000; Bernstein et al., 2005). In contrast, di- or trimethylation of lysine 9 ( H3K9me) and methylation of DNA are usually associated with transcriptionally silence genes and constitutive hetrochromatin (Fahrner et al., 2002; Lehnertz et al., 2003). For plants it has been shown that the majority of H3K9 is mono- or dimethylated (Waterborg, 1990; Jackson et al., 2004).

Currently, it is recognized that both DNA methylation and histone H3K9 methylation are associated with transcriptional silencing. These two modifications share a conspicuous interdependence and are equally important for heterochromatin formation. Two independent genetic screens were performed in the flowering plant Arabidopsis for the mutation that derepress transcriptional silencing at normally heavily methylated SUPERMAN and PAI2 loci (Jackson et al., 2002; Malagnac et al., 2002). In both screens, mutations in the *kryptonite* (*kyp*) gene, which encodes a H3K9-specific methylase with a conserved SET domain (protein lysine methyltransferase enzyme), not only abolished histone methylation at the investigated loci, but also reduced DNA methylation at plant-specific CpNpG and CpNpN trinucleotides (N= A, T, or C). Furthermore, one of the few Arabidopsis homologs of HP1 protein LHP1 was shown to interact directly with CHROMOMETHYLASE 3 (CMT3), a DNA methyltransferase that methylates cytosine at CpNpG and CpNpN sites (Lindroth et al., 2001; Jackson et al., 2002). Interestingly, CMT3, as its name suggests, has chromodomain

motif, which is conserved in proteins that bind to methylated lysine. To bind chromatin, CMT3 requires histone H3 tails to be methylated not only at K9 but also K27 (Lindroth et al., 2004). Thus CMT3 can probably act independently of HP1 at certain loci. Collectively, these observations imply that histone H3K methylation is the primary chromatin modification that guides the establishment of transcriptionally silenced chromatin and subsequently recruits DNA methylation to further stabilized heterochromatin structures. Such model not only placed H3K9 methylation upstream of DNA methylation, but is in agreement with the notion that DNA methylation may not be always be the primary mechanism for silencing genes (Bachman et al., 2003).

The Arabidopsis genome is methylated not only in CpNpG and CpNpN sites but also at CpG dinucleotides. The enzymes responsible for CpG methylation are the maintenance DNA methyltransferase MET1 and *de novo* DNA methyltransferase DRM1 and DRM2 (Genger et al., 1999; Tariq and Paszkowski, 2004). CpG methylation of DNA is not affected by the loss of H3K9 methylation in kryptonite mutants (Jackson et al., 2002). On the contrary, maintenance DNA-methyltransferase-null plants met1-3 and met 1-4) clearly show a dramatic loss of H3K9 methylation from heterochromatic centromeric and pericentromeric chromosomal regions (chromocenters), although the overall levels of H3K9 methylation in met1-null plants remain relatively normal (Tariq et al., 2003). Together with CpG methylation, CpNpG and CpNpN methylation are also reduced in met1 mutants. This is most likely a secondary effect that follows the reduction of H3K9 methylation in constitutive heterochromatin. Cytological observations have shown that, although MET1 is essential for plant development, loss of DNA methylation and decrease in H3K9 methylation at chromocenters of met1 have no effect on the structural integrity of heterochromatin (Tariq et al., 2003). Therefore, it has been argued that additional factors may be involved in the maintenance of chromatin structure at these regions.

Taken together, these studies of Arabidopsis demonstrate the existence of self-reinforcing system that contributes to the formation of silenced chromatin,

where DNA methylation of CpG sites by MET1 plays a key role and is clearly upstream of chromatin modifications.

In other organisms, *Neurospora crassa* a filamentous fungus, the amount of methylated DNA within the genome is relatively low level (~ 2% of all cytosines). DNA methylation is introduced into the Neurospora genome by a mechanism that sends signals from histone to DNA and requires the presence of trimethylated K9 of H3 and binding of Hp1 to modified nucleosomes (Jackson et al., 2004). Fungi and plants share some conserved mechanisms of heterochromatin formation, where H3K9 methylation is either completely (Neurospora) or partially (Arabidopsis) responsible for DNA methylation patterns in the genome. However, DNA methylation is essential for the appropriate localization of H3K9 methylation to heterochromatin in plants, but it is completely dispensable in filamentous fungi *Neurospora*. Another difference between these species is that trimethylation of H3K9 is the predominant heterochromatic modification in *Neurospora*, whereas for plants dimethylation of H3K9 has been suggested to play a major role in gene silencing (Jackson et al., 2004).

## 8: The role of chromatin remodeling activities and RNA interference in the formation of heterochromatin

The mechanisms that lead to the establishment of DNA methylation and histone H3K9 methylation and histone H3K9 methylation patterns are largely unknown, but some studies of plants have allowed us to distinguish initiation and maintenance of heterochromatin-induced gene silencing. A phenomenon known as RNA-directed DNA methylation (RdDM), first described in tobacco plants and later demonstrated in other plant species, provides an important mechanistic link between RNA interference, initiation of gene silencing, and DNA methylation (Pelissier et al., 1999).Genetic analysis indicates that the plant de novo DNA methyltransferase DRM1 and DRM2 are involved in the initiation of the silencing process; they introduce DNA methylation at CpG and non-CpG sites (Cao and Jacobsen, 2002; Chan et al., 2004). Furthermore, mutations in various components of RNAi pathway in Arabidopsis, such as *dicer-like 3* (*dcr3*),

*argonaute 4* (*ago4*)*, RNA-dependent RNA polymerase2* (*rdr2*)*, and silencing-defective 4* (*sde4*), abolish RdDM at various loci, suggesting a direct connection between RNAi and gene silencing by *de novo* DNA methylation (Chan et al., 2004; Tariq and Paszkowski, 2004).

Notably, DRM enzymes are not involved in the maintenance of gene silencing in plants after dsRNA has been degraded (Cao and Jacobsen, 2002). Maintenance of the silenced state requires DNA methyltransferase MET1 and a chromatin remodeling gene known as DDM1 ( deficient in DNA methylation 1) (Morel et al., 2000; Jones et al., 2001). How DDM1 functions is largely unknown, but ddm1-deficient plants lose DNA methylation globally in all sequence contexts, and lose H3K9 methylation from chromocenters and repetitive DNA elements (Gendrel et al., 2002; Lippman et al., 2004). Another SW1/SNF2 family protein, DRD1, with a more specific function in RdDM, has also been identified in *Arabidopsis* (Kanno et al., 2004). Unlike DDM1, DRD1 mutations do not result in a global loss of DNA methylation at target sequences. Thus , in plants, a picture is emerging in which dsRNA, components of the RNAi processing pathways, de novo methylase, and chromatin remodeling activities are all involved in the early steps of heterochromatin formation and initiation of gene silencing. The silenced state is further stabilized by the maintenance CpG DNA methyltransferase MET1, which recruits histone H3K9 methylase (KRYPTONYTE) and non-CpG methylase, such as CHROMOMETHYLASE 3 (Fig. 7).

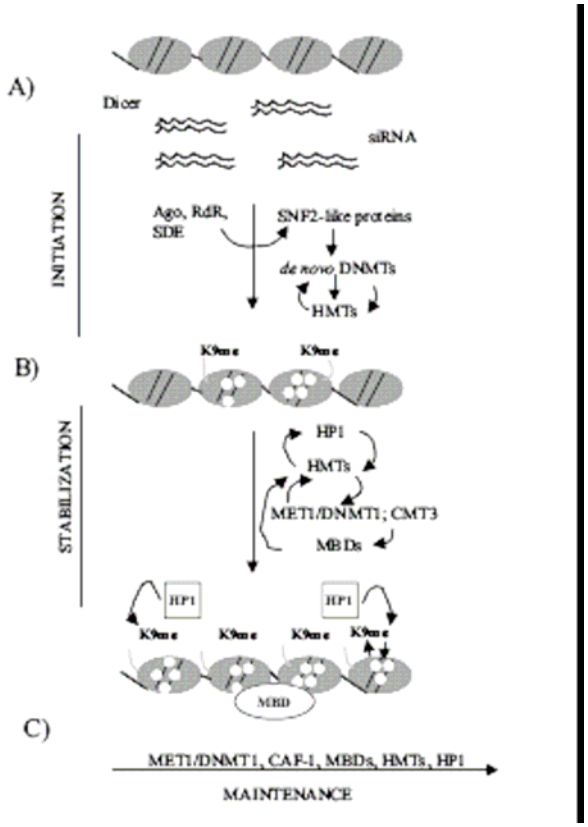


**Fig 7:** Scheme of heterochromatin formation, stabilization and maintenance in plants and mammals. A) During the initial stage of heterochromatin formation, RNA interference pathway, SNF2-like chromatin remodeling activities, de novo DNA methyltrasferase (DNMTs), and histone methylase3 (HMT) introduce silencing epigenetic modifications such as, DNA methylation (white dots on DNA wrapped around the histone core), and histone H3K9 methylation (K9m) into chromatin. DNA methylation and H3K9 methylation mutually reinforce each other in positive feed back loop (small arrows). B) Heterochromatin is further stabilized by the recruitment of maintenance methyltransferase (DNMT1 or MET1), non-CpG methylase (s) (CMT3), HP1 and methyl-CpG binding proteins (MBDs). There is a complex cross talk between various components of DNA methylation and histone methylation pathways (indicated by small arrows). C) Maintenance of heterochromatin through the cooperative action of DNMT1 or MET1, chromatin assembly factor CAF-1, methyl-CpG binding protein (s), histone methylase and most likely HP1.

Further maintenance of heterochromatin structure through cell division is achieved by the cooperative function of maintenance DNA methyltransferase, histone H3K9 methylases, methyl-CpG binding proteins, HP1 proteins, chromatin assembly factor CAF-1 (Shibahara and Stillman, 1999) and methyl-CpG binding protein MBD1(Sarraf and Stancheva, 2004). In the maintenance process, DNA methylation and histone H3K9 methylation mutually reinforce each other to

preserve the integrity of heterochromatic regions and to transmit stable gene silencing through DNA replication (Sarraf and Stancheva, 2004).

## 9: References


**Abdel Ghany, A.G., and Zaki, E.A.** (2002). Cloning and sequencing of an envelope-like gene in Gossypium. Planta **216,** 351-353.

**Almeida, R., and Allshire, R.C.** (2005). RNA silencing and genome regulation. Trends Cell Biol **15,** 251-258.

**Ando, T.** (1996). Destribution of *Petunia axillaris* (Solanaceae) and its new subspecies in Argentina and Bolivia. Acta Phytotax Geobot **47,** 19-30.

**Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, A.J., and Matzke, M.** (2002). RNA-directed DNA methylation in Arabidopsis. Proc Natl Acad Sci U S A **99 Suppl 4,** 16499-16506.

**Aukerman, M.J., and Sakai, H.** (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell **15,** 2730-2741.

**Bachman, K.E., Park, B.H., Rhee, I., Rajagopalan, H., Herman, J.G., Baylin, S.B., Kinzler, K.W., and Vogelstein, B.** (2003). Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. Cancer Cell **3,** 89-95.

**Bailey, L.H.** (1986). Evolution of the *Petunia*. In the Survival of the Unlike, MacMillan**,** 465-472.

**Baulcombe, D.** (2004). RNA silencing in plants. Nature **431,** 356-363.

**Baumberger, N., and Baulcombe, D.C.** (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A **102,** 11928-11933.

**Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., Schreiber, S.L., and Lander, E.S.** (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell **120,** 169-181.

**Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J.** (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature **409,** 363-366.

**Bird, A.** (2002). DNA methylation patterns and epigenetic memory. Genes Dev **16,** 6-21.

**Boeke, J.D., and Corces, V.G.** (1989). Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol **43,** 403-434.

**Boeke, J.D., and Devine, S.E.** (1998). Yeast retrotransposons: finding a nice quiet neighborhood. Cell **93,** 1087-1089.

**Boeke, J.D., Eichinger, D., Castrillon, D., and Fink, G.R.** (1988). The Saccharomyces cerevisiae genome contains functional and nonfunctional copies of transposon Ty1. Mol Cell Biol **8,** 1432-1442.

**Boeke, J.D., Eickbush, T.H., Sandmeyer, S.B., and Voytas, D.F.** (2000a). Metaviridae. vanRegenmortel, M. H. V., Fauquet C.M., Bishop, D. H. L. Carsten, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., and Wickner, R. B. eds. Virus taxonomy: seventh report of international committee on taxonomy of viruses. Academic press, Newyork.**,** 359-367.

**Boeke, J.D., Eickbush, T.H., Sandmeyer, S.B., and Voytas, D.F.** (2000b). Pseudoviridae. pp 349-357in vanRegenmortel, M. H. V., Fauquet C.M., Bishop, D. H. L. Carsten, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., and Wickner, R. B. eds. Virus taxonomy: seventh report of international committee on taxonomy of viruses. Academic press, Newyork.

**Bonneville, J.M., and Hohn, T.** (1993). A reverse transcriptase for cauliflower mosaic virus. The state of art, 1991. In Reverse transcriptase, N. Skalka and S. Goff, eds. Cold Spring Harbor Laboratory Press**,** 357-390.

**Boris-Lawrie, K., and Temin, H.M.** (1994). The retroviral vector. Replication cycle and safety considerations for retrovirus-mediated gene therapy. Ann N Y Acad Sci **716,** 59-70; discussion 71.

**Brown, P.O.** (1990). Integration of retroviral DNA. Curr Top Microbiol Immunol **157,** 19-48.

**Buratowski, S., and Moazed, D.** (2005). Gene regulation: expression and silencing coupled. Nature **435,** 1174-1175.

**Cao, X., and Jacobsen, S.E.** (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci U S A **99 Suppl 4,** 16491-16498.

**Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E.** (2003). Role of the DRM and CMT3

methyltransferases in RNA-directed DNA methylation. Curr Biol **13,** 2212-2217.

**Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E.** (2004). RNA silencing genes control de novo DNA methylation. Science **303,** 1336.

**Chavanne, F., Zhang, D.X., Liaud, M.F., and Cerff, R.** (1998). Structure and evolution of Cyclops: a novel giant retrotransposon of the Ty3/Gypsy family highly amplified in pea and other legume species. Plant Mol Biol **37,** 363-375.

**Coffin, J.M., Hughes, S.H., and Varmus, H.E.** (1997). Retroviruses. Cold Spring Harbor Laboratory Press.

**Cogoni, C., and Macino, G.** (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature **399,** 166-169.

**Cogoni, C., Irelan, J.T., Schumacher, M., Schmidhauser, T.J., Selker, E.U., and Macino, G.** (1996). Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. Embo J **15,** 3153-3163.

**Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D.C.** (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell **101,** 543-553.

**Devine, S.E., and Boeke, J.D.** (1996). Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. Genes Dev **10,** 620-633.

**Eickbush, T.H.** (2002). Repair by retrotransposition. Nat Genet **31,** 126-127.

**Eigel, A., and Feldmann, H.** (1982). Ty1 and delta elements occur adjacent to several tRNA genes in yeast. Embo J **1,** 1245-1250.

**Fagard, M., and Vaucheret, H.** (2000a). (TRANS)GENE SILENCING IN PLANTS: How Many Mechanisms? Annu Rev Plant Physiol Plant Mol Biol **51,** 167-194.

**Fagard, M., and Vaucheret, H.** (2000b). Systemic silencing signal(s). Plant Mol Biol **43,** 285-293.

**Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H.** (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-

transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A **97,** 11650-11654.

**Fahrner, J.A., Eguchi, S., Herman, J.G., and Baylin, S.B.** (2002). Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. Cancer Res **62,** 7213-7218.

**Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C.** (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature **391,** 806-811.

**Fitchen, J.H., and Beachy, R.N.** (1993). Genetically engineered protection against viruses in transgenic plants. Annu Rev Microbiol **47,** 739-763.

**Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., and Nakao, M.** (1999). Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. Mol Cell Biol **19,** 6415-6426.

**Ganem, D., and Varmus, H.E.** (1987). The molecular biology of the hepatitis B viruses. Annu Rev Biochem **56,** 651-693.

**Geering, A.D., Olszewski, N.E., Harper, G., Lockhart, B.E., Hull, R., and Thomas, J.E.** (2005). Banana contains a diverse array of endogenous badnaviruses. J Gen Virol **86,** 511-520.

**Geering, A.D.W., Olszewski, N.E., Dahal, G., Thomas, J.E., and Lockhart, B.E.L.** (2001). Analysis of the distribution and structure of integrated Banana strek virus DNA in a range of Musa cultivars. Mol Plant Pathol **2,** 207-213.

**Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A.** (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science **297,** 1871-1873.

**Genger, R.K., Kovac, K.A., Dennis, E.S., Peacock, W.J., and Finnegan, E.J.** (1999). Multiple DNA methyltransferase genes in Arabidopsis thaliana. Plant Mol Biol **41,** 269-278.

**Gerats, A.G., Huits, H., Vrijlandt, E., Marana, C., Souer, E., and Beld, M.** (1990). Molecular characterization of a nonautonomous transposable element (dTph1) of petunia. Plant Cell **2,** 1121-1128.

**Gerats, T., and Vandenbussche, M.** (2005). A model system for comparative research: Petunia. Trends Plant Sci **10,** 251-256.

**Goodall, G.J., and Filipowicz, W.** (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58,** 473-483.

**Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G.** (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature **386,** 44-51.

**Grandbastien, M.A.** (1998a). Activation of plant retrotransposons under stress conditions. Trends Cell Biol **3,** 181-187.

**Grandbastien, M.A.** (1998b). Activation of plant retrotransposons under stress conditions. Trends Plant Sci **3,** 181-187.

**Grandbastien, M.A.** (2004). [Stress activation and genomic impact of plant retrotransposons]. J Soc Biol **198,** 425-432.

**Gregor, W., Mette, M.F., Staginnus, C., Matzke, M.A., and Matzke, A.J.** (2004). A distinct endogenous pararetrovirus family in Nicotiana tomentosiformis, a diploid progenitor of polyploid tobacco. Plant Physiol **134,** 1191-1199.

**Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J.** (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature **404,** 293-296.

**Hansen, C., and Heslop-Harrison, J.S.** (2004). Sequences and phylogenies of plant pararetroviruses, viruses and transposable elements. Advances in botanical research **41,** 165-193.

**Harper, G., and Hull, R.** (1998). Cloning and sequence analysis of banana streak virus DNA. Virus Genes **17,** 271-278.

**Harper, G., Dahal, G., Thottappilly, G., and Hull, R.** (1999a). Detection of episomal banana streak badnavirus by IC-PCR. J Virol Methods **79,** 1-8.

**Harper, G., Osuji, J.O., Heslop-Harrison, J.S., and Hull, R.** (1999b). Integration of banana streak badnavirus into the Musa genome: molecular and cytogenetic evidence. Virology **255,** 207-213.

**Harper, G., Hull, R., Lockhart, B., and Olszewski, N.** (2002). Viral sequences integrated into plant genomes. Annu Rev Phytopathol **40,** 119-136.

**Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C.** (2005). RNA polymerase IV directs silencing of endogenous DNA. Science **308,** 118-120.

**Heslop-Harrison, J.S.** (2000). Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. Plant Cell **12,** 617-636.

**Hughes, J.F., and Coffin, J.M.** (2001). Evidence for genomic rearrangements mediated by human endogenous retroviruses during primate evolution. Nat Genet **29,** 487-489.

**Hull, R.** (1999). Classification of reverse transcribing elements: a discussion document. Arch Virol **144,** 209-213; discussion 213-204.

**Hull, R.** (2001). Classifying reverse transcribing elements: a proposal and a challenge to the ICTV. International Committee on Taxonomy of Viruses. Arch Virol **146,** 2255-2261.

**Hull, R., and Covey, S.N.** (1995). Retroelements: propagation and adaptation. Virus Genes **11,** 105-118.

**Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E.** (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416,** 556-560.

**Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E.** (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma **112,** 308-315.

**Jakowitsch, J., Mette, M.F., van Der Winden, J., Matzke, M.A., and Matzke, A.J.** (1999). Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proc Natl Acad Sci U S A **96,** 13241-13246.

**Jenuwein, T., and Allis, C.D.** (2001). Translating the histone code. Science **293,** 1074-1080.

**Ji, H., Moore, D.P., Blomberg, M.A., Braiterman, L.T., Voytas, D.F., Natsoulis, G., and Boeke, J.D.** (1993). Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. Cell **73,** 1007-1018.

**Johansen, L.K., and Carrington, J.C.** (2001). Silencing on the spot. Induction and suppression of RNA silencing in the Agrobacterium-mediated transient expression system. Plant Physiol **126,** 930-938.

**Johnson, L., Cao, X., and Jacobsen, S.** (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Curr Biol **12,** 1360-1367.

**Jones, L., Ratcliff, F., and Baulcombe, D.C.** (2001). RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Curr Biol **11,** 747-757.

**Kalantidis, K., Psaradakis, S., Tabler, M., and Tsagris, M.** (2002). The occurrence of CMV-specific short Rnas in transgenic tobacco expressing virus-derived double-stranded RNA is indicative of resistance to the virus. Mol Plant Microbe Interact **15,** 826-833.

**Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., and Matzke, A.J.** (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr Biol **14,** 801-805.

**Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J.** (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet **37,** 761-765.

**Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y.** (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science **309,** 467-469.

**Kawasaki, H., and Taira, K.** (2005). Transcriptional gene silencing by short interfering RNAs. Curr Opin Mol Ther **7,** 125-131.

**Ketting, R.F., and Plasterk, R.H.** (2000). A genetic link between co-suppression and RNA interference in C. elegans. Nature **404,** 296-298.

**Kim, A., Terzian, C., Santamaria, P., Pelisson, A., Purd'homme, N., and Bucheton, A.** (1994). Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of Drosophila melanogaster. Proc Natl Acad Sci U S A **91,** 1285-1289.

**Kirchner, J., Connolly, C.M., and Sandmeyer, S.B.** (1995). Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. Science **267,** 1488-1491.

**Kouzarides, T.** (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev **12,** 198-209.

**Kumar, A., and Bennetzen, J.L.** (1999). Plant retrotransposons. Annu Rev Genet **33,** 479-532.

**Kumar, A., and Bennetzen, J.L.** (2000). Retrotransposons: central players in the structure, evolution and function of plant genomes. Trends Plant Sci **5,** 509-510.

**Kunii, M., Kanda, M., Nagano, H., Uyeda, I., Kishima, Y., and Sano, Y.** (2004). Reconstruction of putative DNA virus from endogenous rice tungro bacilliform virus-like sequences in the rice genome: implications for integration and evolution. BMC Genomics **5,** 80.

**Lachner, M., and Jenuwein, T.** (2002). The many faces of histone lysine methylation. Curr Opin Cell Biol **14,** 286-298.

**Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T.** (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature **410,** 116-120.

**Laten, H.M., Majumdar, A., and Gaucher, E.A.** (1998). SIRE-1, a copia/Ty1-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc Natl Acad Sci U S A **95,** 6897-6902.

**Leclerc, D., Chapdelaine, Y., and Hohn, T.** (1999). Nuclear targeting of the cauliflower mosaic virus coat protein. J Virol **73,** 553-560.

**Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.** (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol **13,** 1192-1200.

**Lesemann, D.-E., and Casper, R.** (1973). Electron microscopy of petunia vein clearing virus, an isometric plant virus associated with specific inclusions in petunia cells. Phytopathology **63,** 1118-1124.

**Li, L.C., and Dahiya, R.** (2002). MethPrimer: designing primers for methylation PCRs. Bioinformatics **18,** 1427-1431.

**Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E.** (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science **292,** 2077-2080.

**Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh, S., and Jacobsen, S.E.** (2004). Dual

histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. Embo J **23,** 4286-4296.

**Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R.** (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol **1,** E67.

**Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V., and Martienssen, R.** (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature **430,** 471-476.

**Lockhart, B.E., Menke, J., Dahal, G., and Olszewski, N.E.** (2000). Characterization and genomic analysis of tobacco vein clearing virus, a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. J Gen Virol **81,** 1579-1585.

**Lockhart, B.E.L.** (1986). Purification and serology of a bacilliform virus associated with banana streak disease. Phytopathology **76,** 995-999.

**Lockhart, B.E.L., and Olszewski, N.E.** (1993). Serological and genomic heterogeneity of banana streak badnavirus: implications for virus detection in Musa germplasm. In: Ganry, J. (Ed.),. Breeging banana and plantain for resistance to diseases and pets. CIRAD/INIBAP, Montpellier, France**,** 105-113.

**Lockhart, B.E.L., and Lesemann, D.-E.** (1997). Occurrence of petunia vein clearing virus in the U. S. A. Plant Dis **82,** 262.

**Makeyev, E.V., and Bamford, D.H.** (2002). Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. Mol Cell **10,** 1417-1427.

**Malagnac, F., Bartee, L., and Bender, J.** (2002). An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. Embo J **21,** 6842-6852.

**Malik, H.S., and Eickbush, T.H.** (1999). Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. J Virol **73,** 5186-5190.

**Malik, H.S., Henikoff, S., and Eickbush, T.H.** (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res **10,** 1307-1318.

**Marin, I., and Llorens, C.** (2000). Ty3/Gypsy retrotransposons: description of new Arabidopsis thaliana elements and evolutionary perspectives derived from comparative genomic data. Mol Biol Evol **17,** 1040-1049.

**Martienssen, R.A.** (2003). Maintenance of heterochromatin by RNA interference of tandem repeats. Nat Genet **35,** 213-214.

**Martienssen, R.A., and Colot, V.** (2001). DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science **293,** 1070-1074.

**Matzke, M., Matzke, A.J., and Kooter, J.M.** (2001). RNA: guiding gene silencing. Science **293,** 1080-1083.

**Matzke, M., Aufsatz, W., Kanno, T., Daxinger, L., Papp, I., Mette, M.F., and Matzke, A.J.** (2004). Genetic analysis of RNA-mediated transcriptional gene silencing. Biochim Biophys Acta **1677,** 129-141.

**Matzke, M.A., and Birchler, J.A.** (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet **6,** 24-35.

**Matzke, M.A., Aufsatz, W., Kanno, T., Mette, M.F., and Matzke, A.J.** (2002). Homology-dependent gene silencing and host defense in plants. Adv Genet **46,** 235-275.

**Meehan, R.R., and Stancheva, I.** (2001). DNA methylation and control of gene expression in vertebrate development. Essays Biochem **37,** 59-70.

**Menissier, J., de Murcia, G., Lebeurier, G., and Hirth, L.** (1983). Electron microscopic studies of the different topological forms of the cauliflower mosaic virus DNA: knotted encapsidated DNA and nuclear minichromosome. Embo J **2,** 1067-1071.

**Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J.** (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. Embo J **19,** 5194-5201.

**Mette, M.F., Kanno, T., Aufsatz, W., Jakowitsch, J., van der Winden, J., Matzke, M.A., and Matzke, A.J.** (2002). Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. Embo J **21,** 461-469.

**Metzlaff, M., O'Dell, M., Cluster, P.D., and Flavell, R.B.** (1997). RNA-mediated RNA degradation and chalcone synthase A silencing in petunia. Cell **88,** 845-854.

**Mewes, H.W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S.G., Pfeiffer, F., and Zollner, A.** (1997). Overview of the yeast genome. Nature **387,** 7-65.

**Mochizuki, K., and Gorovsky, M.A.** (2004). Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. Genes Dev **18,** 2068-2073.

**Morel, J.B., Mourrain, P., Beclin, C., and Vaucheret, H.** (2000). DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in Arabidopsis. Curr Biol **10,** 1591-1594.

**Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T.A., and Vaucheret, H.** (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell **101,** 533-542.

**Murfett, J., Wang, X.J., Hagen, G., and Guilfoyle, T.J.** (2001). Identification of Arabidopsis histone deacetylase HDA6 mutants that affect transgene expression. Plant Cell **13,** 1047-1061.

**Napoli, C., Lemieux, C., and Jorgensen, R.** (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell **2,** 279-289.

**Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olszewski, N.E., and Lockhart, B.** (1999). Evidence that badnavirus infection in Musa can originate from integrated pararetroviral sequences. Virology **255,** 214-220.

**Onodera, Y., Haag, J.R., Ream, T., Nunes, P.C., Pontes, O., and Pikaard, C.S.** (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell **120,** 613-622.

**Palauqui, J.C., Elmayan, T., Pollien, J.M., and Vaucheret, H.** (1997). Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. Embo J **16,** 4738-4745.

**Park, W., Li, J., Song, R., Messing, J., and Chen, X.** (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol **12,** 1484-1495.

**Pelissier, T., Thalmeir, S., Kempe, D., Sanger, H.L., and Wassenegger, M.** (1999). Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucleic Acids Res **27,** 1625-1634.

**Peterson, C.L., and Laniel, M.A.** (2004). Histones and histone modifications. Curr Biol **14,** R546-551.

**Peterson-Burch, B.D., and Voytas, D.F.** (2002). Genes of the Pseudoviridae (Ty1/copia retrotransposons). Mol Biol Evol **19,** 1832-1845.

**Peterson-Burch, B.D., Wright, D.A., Laten, H.M., and Voytas, D.F.** (2000). Retroviruses in plants? Trends Genet **16,** 151-152.

**Pfeiffer, P., and Hohn, T.** (1983). Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell **33,** 781-789.

**Pooggin, M., Shivaprasad, P.V., Veluthambi, K., and Hohn, T.** (2003). RNAi targeting of DNA virus in plants. Nat Biotechnol **21,** 131-132.

**Potter, S.S., Brorein, W.J., Jr., Dunsmuir, P., and Rubin, G.M.** (1979). Transposition of elements of the 412, copia and 297 dispersed repeated gene families in Drosophila. Cell **17,** 415-427.

**Probst, A.V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., Lawrence, R.J., Pikaard, C.S., Murfett, J., Furner, I., Vaucheret, H., and Scheid, O.M.** (2004). Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. Plant Cell **16,** 1021-1034.

**Prols, F., and Meyer, P.** (1992). The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida. Plant J **2,** 465-475.

**Qi, Y., Denli, A.M., and Hannon, G.J.** (2005). Biochemical Specialization within Arabidopsis RNA Silencing Pathways. Mol Cell **19,** 421-428.

**Richards, E.J.** (1997). DNA methylation and plant development. Trends Genet **13,** 319-323.

**Richards, E.J., and Elgin, S.C.** (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108,** 489-500.

**Richert-Poggeler, K.R., and Shepherd, R.J.** (1997). Petunia vein-clearing virus: a plant pararetrovirus with the core sequences for an integrase function. Virology **236,** 137-146.

**Richert-Poggeler, K.R., Noreen, F., Schwarzacher, T., Harper, G., and Hohn, T.** (2003). Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. Embo J **22,** 4836-4845.

**Rothnie, H.M., Chapdelaine, Y., and Hohn, T.** (1994). Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv Virus Res **44,** 1-67.

**Sanfacon, H., and Hohn, T.** (1990). Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. Nature **346,** 81-84.

**Sarraf, S.A., and Stancheva, I.** (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell **15,** 595-605.

**Schramke, V., and Allshire, R.** (2003). Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science **301,** 1069-1074.

**Schramke, V., Sheedy, D.M., Denli, A.M., Bonila, C., Ekwall, K., Hannon, G.J., and Allshire, R.C.** (2005). RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription. Nature **435,** 1275-1279.

**Shibahara, K., and Stillman, B.** (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell **96,** 575-585.

**Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A.** (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell **107,** 465-476.

**Sink, K.C.** (1984). Taxonomy. In Monographs on Theoretical and applied Genetics: petunia, Springer Verlag **9**.

**Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M.** (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol **10,** 169-178.

**Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M.** (2000). Total silencing by intron-spliced hairpin RNAs. Nature **407,** 319-320.

**Song, S.U., Gerasimova, T., Kurkulos, M., Boeke, J.D., and Corces, V.G.** (1994). An env-like protein encoded by a Drosophila retroelement: evidence that gypsy is an infectious retrovirus. Genes Dev **8,** 2046-2057.

**Stoye, J.P.** (2001). Endogenous retroviruses: still active after all these years? Curr Biol **11,** R914-916.

**Suoniemi, A., Tanskanen, J., and Schulman, A.H.** (1998). Gypsy-like retrotransposons are widespread in the plant kingdom. Plant J **13,** 699-705.

**Tang, W., and Leisner, S.** (1998). Methylation of nonintegrated multiple copy DNA in plants. Biochem Biophys Res Commun **245,** 403-406.

**Tanne, E., and Sela, I.** (2005). Occurrence of a DNA sequence of a non-retro RNA virus in a host plant genome and its expression: evidence for recombination between viral and host RNAs. Virology **332,** 614-622.

**Tariq, M., and Paszkowski, J.** (2004). DNA and histone methylation in plants. Trends Genet **20,** 244-251.

**Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y., and Paszkowski, J.** (2003). Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. Proc Natl Acad Sci U S A **100,** 8823-8827.

**Temin, H.M.** (1970). Malignant transformation of cells by viruses. Perspect Biol Med **14,** 11-26.

**Tenllado, F., Llave, C., and Diaz-Ruiz, J.R.** (2004). RNA interference as a new biotechnological tool for the control of virus diseases in plants. Virus Res **102,** 85-96.

**Turner, B.M.** (2000). Histone acetylation and an epigenetic code. Bioessays **22,** 836-845.

**Vaistij, F.E., Jones, L., and Baulcombe, D.C.** (2002). Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. Plant Cell **14,** 857-867.

**Vaughn, M.W., and Martienssen, R.A.** (2005). Finding the right template: RNA Pol IV, a plant-specific RNA polymerase. Mol Cell **17,** 754-756.

**Vicient, C.M., Kalendar, R., and Schulman, A.H.** (2001). Envelope-class retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. Genome Res **11,** 2041-2049.

**Vicient, C.M., Kalendar, R., and Schulman, A.H.** (2005). Variability, Recombination, and Mosaic Evolution of the Barley BARE-1 Retrotransposon. J Mol Evol.

**Vogt, P.K.** (1997). Historical introduction to the general properties of retroviruses, in Coffin J. M., Hughes S.H., and Varmus H. E., (eds),. Retroviruses, CSHL Press , New York**,** 1-25.

**Volpe, T., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C.** (2003). RNA interference is required for normal centromere function in fission yeast. Chromosome Res **11,** 137-146.

**Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A.** (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science **297,** 1833-1837.

**Walbot, V.** (1999). UV-B damage amplified by transposons in maize. Nature **397,** 398-399.

**Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L.** (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell **76,** 567-576.

**Waterborg, J.H.** (1990). Sequence analysis of acetylation and methylation in two histone H3 variants of alfalfa. J Biol Chem **265,** 17157-17161.

**Waterhouse, P.M., Graham, M.W., and Wang, M.B.** (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc Natl Acad Sci U S A **95,** 13959-13964.

**Wijsman, H.J.W.** (1982). On the interrelationship of certain speices of petunia. Acta. Bot. Neerl **31,** 477-490.

**Wright, D.A., and Voytas, D.F.** (1998). Potential retroviruses in plants: Tat1 is related to a group of Arabidopsis thaliana Ty3/gypsy retrotransposons that encode envelope-like proteins. Genetics **149,** 703-715.

**Xie, Z., Fan, B., Chen, C., and Chen, Z.** (2001). An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. Proc Natl Acad Sci U S A **98,** 6516-6521.

**Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C.** (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol **2,** E104.

**Xiong, Y., and Eickbush, T.H.** (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. Embo J **9,** 3353-3362.

**Zeidan, M., Sikron, N., Cohen, J., and Gera, A.** (2001). Improved detection of petunia vein clearing caulimovirus. HortScience **35**.

# Chapter 2

## Different heterochromatic traits of endogenous *Petunia vein clearing virus* in two distinct *Petunia hybrida* lines

Faiza Noreen, Rashid Akbergenov, Thomas Hohn* and Katja R. Richert-Pöggeler*

Friedrich Miescher Institute, Basel and Botanical Institute of the University of Basel, Switzerland.

*Corresponding authors

(Submitted in The plant cell)

## 1: Abstract

Multiple copies of *Petunia vein clearing virus-1* (PVCV-1) sequences are integrated into the pericentromeric regions of the chromosomes of *Petunia hybrida* as “endogenous PVCV-1” (ePVCV-1), forming part of the silent heterochromatin. Using RT-PCR, we detected higher levels of ePVCV-1 transcripts in symptomatic plant tissue than in non-symptomatic tissue, indicative of PVCV infection derived from induction of ePVCV-1 sequences. To correlate ePVCV-1 transcript levels with epigenetic markers, we examined DNA- and histone-methylation and siRNA production in cultivars W138 and Rdc. Similar levels of ePVCV-1 cytosine methylation were observed in both hosts. In contrast, significant differences in histone modification were observed: in Rdc, ePVCV-1 sequences were mostly associated with histone 3 methylated at lysine 9 (H3mK9), a flag for heterochromatin, while in W138 they were equally associated with H3mK9 and H3mK4, the latter being a marker for euchromatin. Interestingly, the H3mK4/H3mK9 ratio was relatively high in the promoter region in both cultivars. In symptomatic leaves, virus-specific siRNAs 21, 22 and 24 nt in size were detected. Our results suggest that a defect in the silencing pathway in W138 could explain the higher activity of ePVCV-1 sequences in this cultivar.

## 2: Introduction

Integrated viral sequences are familiar components of bacterial and animal genomes. For instance, about 8% of the human genome is derived from retroviruses (Hughes and Coffin, 2001). In contrast, plant genomes have generally been thought to be free of viral sequences. This concept has changed during the last few years as DNA of both geminiviruses and pararetroviruses has been found integrated into the plant genome (Harper et al., 2002; Kunii et al., 2004). Furthermore, apparently reverse transcribed sequences derived from an RNA virus (potyvirus) have been isolated from grapevine (Tanne and Sela, 2005).

The invasion of plant genomes by mobile elements of viral origin, and the risk of their reactivation, deserves the attention of host control mechanisms, which are related to the epigenetic control of gene activity. Such control is exerted by chromatin condensation of the wider heterochromatin, encompassing pericentromeric and telomeric regions (Richards, 1997) or by dispersed heterochromatic polycomb-type regulated domains (Goodrich et al., 1997). Small interfering RNAs (siRNAs) serve as signaling molecules, triggering *de novo* methylation of cytosines and modification of the associated histones, exemplified by methylation of lysine 9 and/or 27 of histone 3 (H3mK9 and H3mK27; Matzke and Birchler 2005; Jackson et al., 2004; Lindroth et al., 2004; Mathieu et al., 2005). During replication, methylation in symmetrical context (CG, CNG) is maintained by the action of maintenance methylases. Further factors involved are the chromatin remodeling ATPase DDM1 (Decrease in DNA Methylation1; Hirochika et al., 2000; Miura et al., 2001) and the putative chromatin remodeling factor MOM1 that, in contrast to DDM1, regulates transcriptional gene silencing (TGS) independently of DNA methylation (Amedeo et al., 2000).

Plant pararetroviruses consist of open circular DNA packaged in icosahedral or bacilliform capsids. The different genera of plant pararetroviruses can be distinguished by genome arrangement and expression strategies (Hohn and Richert-Pöggeler, 2005). These viral genomes exhibit different degrees of gene fusion, with *Petunia vein clearing virus* (PVCV) at one extreme, with all its

genes fused into a single open reading frame (Fig. 1a). The genome manifests itself as a minichromosome from which terminally redundant RNA, resembling retrovirus RNA, is produced. However, reverse transcription of this RNA does not lead to terminally redundant DNA, and the genome lacks an active integrase. Consequently, unlike in retroviruses, integration of plant pararetroviruses cannot be an obligatory part of replication; instead, the genome establishes itself as a circular non-redundant minichromosome. The rare integrated plant pararetroviruses are usually defective and cannot be induced, with the exception of a few examples such as ePVCV-1 (Richert-Pöggeler et al., 2003), *Tobacco vein clearing virus* (TVCV, Lockhart et al., 2000) and *Banana streak virus* strain "Obino l'Ewai" (BSOEV; Harper et al., 1999; Ndowora et al., 1999).

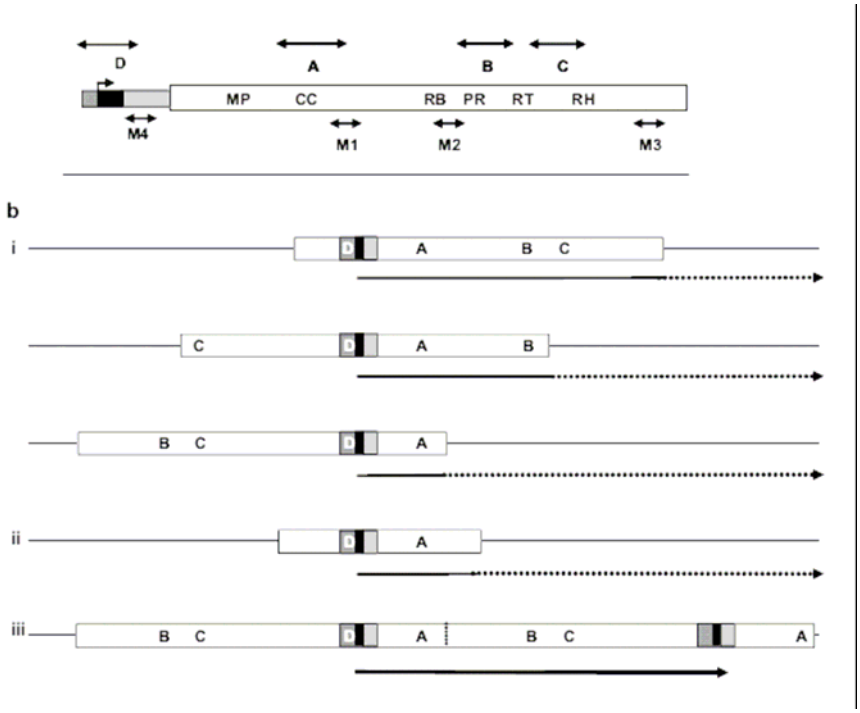


**Fig 1.** a) Schematic map of PVCV, which consists of a single ORF (large rectangle) and an intergenic region (small rectangle) containing a promoter and other signals (TATA box, polyA). Plant pararetrovirus consensus motifs (MP, movement protein domain; CC, coiled-coil domain; RB, RNA-binding domain; PR, protease; RT, reverse transcriptase; RH, RNAse H) are indicated. The positions of primers used for RT-PCR and ChIP experiments (A-D), and bisulfite sequencing (M1-M4), are shown above and below the map, respectively. The bent arrow indicates the transcription start site. b) Model for RNAs transcribed under the control of the PVCV promoter from integrated PVCV DNA. Bold arrows represent PVCV-specific RNA sequences; dotted sections represent readthrough into surrounding host sequences. The triple box represents the quasi-LTR (QTR: dark grey: U3 with promoter; black: R with polyadenylation signal; light grey: U5). Note that the promoter region is transcribed only from tandem arrays. i) A series of full-

length integrates originating from different sites within the PVCV DNA, ii) an integrated fragment of PVCV DNA, iii) tandem integrants of PVCV DNA.

Integrated forms of ePVCV-1 are found in the pericentromeric regions of *Petunia hybrida* chromosomes (Richert-Pöggeler et al., 2003; K. Richert-Pöggeler and T. Schwarzacher, unpublished). Like retroelements (Walbot, 1999; Grandbastien, 2004), ePVCV-1 is induced under stress conditions (Richert-Pöggeler et al., 2003), leading to systemic infection as evidenced by vein clearing symptoms. Induction could occur via homologous recombination between repetitive sequences, as was postulated for BSOEV (Ndowora et al., 1999) or via direct transcription from nested full-length integrants (Richert-Pöggeler et al., 2003).

We selected the original host plant of PVCV -- *P. hybrida* line "Himmelsröschen" (also known as "rose du ciel" or Rdc; Lesemann and Casper, 1973) -- and line W138, which is well known due to its DNA transposon *dTph1*, which is widely applied for mutagenesis (Gerats, 1990; Van Houwelingen et al., 1998; Gerats and Vandenbussche, 2005). *dTph1* is activated by "activator 1" (ACT1*)* by an as yet unknown mechanism (Stuurman and Kuhlemeier, 2005). *dTph1* exists also in Rdc, although with a reduced copy number (10-30 compared to up to 200 in W138), and is inactive, probably due to the lack of an appropriate activator. In this study, we compared the extent and the similarity of epigenetic control of ePVCV-1 and *dTph1* in these two *P. hybrida* lines. We could show that histone modification, rather than DNA methylation, correlates with lower levels of transcription. Based on our results, we propose a threshold model for activation of ePVCV-1 in petunia resulting in systemic infection, and control via the host RNA silencing machinery.

# 3:Materials and Methods

## 3.1: Plant material and DNA extraction

Greenhouse plants were grown under 12-14 h light at 22-25°C. For tissue culture, petunia seeds were surface sterilized. Plants were grown under 16 h light at 25°C and 60% humidity. Genomic plant DNA was extracted using the nucleon

*PhytoPure* extraction kit (Amersham Life Sciences, Little Chalfont, UK). Seeds of *P. hybrida* cv. Rdc were purchased from NL Chrestensen (Erfurt, Germany). Seeds of W138 and *P. parodii*, originally obtained by E. Sour (University of Amsterdam) and J.-P. Renou (INRA-CR Angers), respectively, were amplified in the greenhouse.

### 3.2: Bisulfite –PCR Methylation Analysis

Bisulfite treatment was performed using an EZ DNA Methylation kit (ZYMO Research, Orange, CA) according to the manufacturer's protocol, using 1 µg of genomic DNA. A 2 µl aliquot of each bisulfite–modified DNA sample (200 ng) was subjected to PCR. Primers M-1, -2, -3, and –4 were designed using the MethPrimer program (Li and Dahiya, 2002) towards virus-specific sequences, but with cytosines replacing thymines for the forward primers and guanines replacing adenines for the reverse primers (Table 1). These primer pairs synthesize fragments from bisulfite-treated DNA templates but not from untreated DNA templates. Accordingly, primer pairs containing the original virus sequence (2233-L/2424-R, 3694-L/3998-R, 6193-L/6456-R, see Table 1) produced a PCR fragment only when untreated DNA was present but not when bisulfite-treated DNA was used as template.

For methylation analyses, the plus (transcribed) strands of the PCR products obtained from bisulfite treated DNA were sequenced using the forward primers.

### 3.3: Reverse transcription (RT) – PCR

Total RNA from *P. hybrida* (W138 or Rdc) non-symptomatic (1 and 2 months old) and symptomatic (4 months and older) plants was isolated with an RNeasy Mini Kit (Qiagen, Hilden, Germany) and reverse transcribed using Ready-To-Go You-Prime First-Strand Beads (Amersham) according to the manufacturer's instructions. Synthesis of cDNA was primed with an oligo(dT) primer (Roche, Mannheim, Germany) as follows: 3 µg of extracted RNA, resuspended in RNase free water, was denatured for 10 min at 65°C and placed on ice for 2 min. The RNA solution was transferred to the tube of "First strand

reaction mix beads" and, after addition of primer, incubated for 1 hour at 37°C. A series of 3 dilutions (1:1, 1:20, 1:50) of generated cDNA was prepared and amplified with primers corresponding to the endogenous gene encoding the ribosomal protein large subunit (rpl2) as an internal control (see Table 1). The 1:50 dilution was used thereafter in PCR amplification of both virus-specific and endogenous sequences. Four PCR primer sets were applied in the methylation studies as well as in ChIP experiments (see Table 1). PCR was performed in a 50 µl volume and conditions were as follows: initial denaturation at 94°C for 2 min, 35 cycles of (94°C for 45 s, 50°C for 45 s, 72°C for 1min) and a final step at 72°C for 10 min. For the endogenous sequences the primer annealing temperature was 58°C. A 25 µl aliquot of the amplified DNA was visualized on 1.3% agarose gel stained with ethidium bromide. Control PCR reactions to monitor DNA contamination of the extracted RNA template did not produce a PCR product.

**3.4: Chromatin Immunoprecipitation (ChIP)**

ChIP was carried out as described by Johnson et al. (2002) and Gendrel et al (2002) using 4-week-old soil-grown plants and histone H3 anti-dimethyl lysine-9 or anti-dimethyl lysine-4 antibodies (Upstate Technologies, Lake Placid, NY). Precipitated DNA was resuspended in 40 µl of TE (10 mM Tris, 1 mM EDTA, pH 8). An equal amount of chromatin was mock precipitated without antibody, while a small aliquot of sonicated chromatin was reverse cross-linked, resuspended in 40 µl of TE, diluted, and used as total input DNA control. PCRs were performed in 25 µl with 1 µl of immunoprecipitated DNA. PCR conditions were as follows: 94°C for 2 min, 40 cycles (94°C for 45 s, 50°C for 30 s, 72°C for 1 min), 72°C for 10min. The amplified DNA was visualized on 2% agarose gels stained with ethidium bromide. For control samples (*dTph1* and ribosomal protein large subunit, rpl2) amplification was performed for 30 cycles. In all cases, mock precipitations with no antibody yielded little or no product.

### 3.5: Analysis of siRNA

Total RNA was isolated using Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. In the case of enriched samples, total RNA was fractionated using a Mini RNeasy kit (Qiagen) and the RNA cleanup protocol. 30 µg RNA was re-suspended in 10 µl loading buffer (95% formamide, 20 mM EDTA pH 8.0, 0.05% bromophenol blue, 0.05% xylene cyanol), heated at 95°C for 2 min and loaded on a 15% polyacrylamide gel (a 19:1 ratio of acrylamide to bis-acrylamide, 8M urea). The gel was run using an SE 600 electrophoresis apparatus (Hoefer, San Francisco, CA) at 300V for 4 h and the RNA was then transferred to Hybond N+ membrane (Amersham) by electroblotting in 1xTBE buffer at 10 V overnight. Blot hybridization was performed at 35°C for 14 to 24 h in UltraHyb-oligo buffer (Ambion) using, as a probe, one or several short DNA oligonucleotides (coding region: 2233-L 3694-L, 2233-as and 3694-as; intergenic region: 6798-s and 6798-as; DNA transposon: *dTph1*-s, *dTph1*-2s, *dTph1*-as and *dTph1*-2as: Table 1) end-labeled with $\gamma$–$^{32}$P by polynucleotide kinase (Roche) and purified through MicroSpin™ G-25 columns (Amersham) according to the manufacturer's recommendations. The blot was washed twice with 2x SSC, 0.5% SDS for 30 min at 35°C. The signal was detected after 1-3 days exposure to a phosphor screen using a Molecular Imager (Bio-Rad, Richmond, CA). For repeated hybridization the membrane was stripped with 0.5x SSC, 0.5% SDS for 30 min at 80°C and then with 0.1x SSC, 0.5% SDS for 30 min at 80°C.

The RNA size marker was prepared using T7 RNA polymerase (Promega) and a pair of DNA oligonucleotides as a template – the T7 promoter oligo 5'-TAATACGACTCACTATAG-3' and 5'-ACGGTTGGCCCCTTGGTTTCCCTATAGTGAGTCGTATTA-3' – according to the Promega protocol in the presence of [α-$^{32}$P]rUTP (Hartman, Germany) and then purified through the G-25 columns (Amersham).

# 4:Results

## 4.1: Low-level transcription from silent ePVCV-1

We can distinguish two states of *P. hybrida* / ePVCV-1 interaction. In the non-symptomatic state, the plant is free of viral symptoms and neither episomal

| Name | Orientation | Sequence | ann.temp (˚C ) | nt. position in PVCV | Fragment size |
|---|---|---|---|---|---|
| For Bisulfite sequencing | | | | | |
| M1-L | Forward | TTT TAG TTT AGT AGT TAT TTT GGT GAT TAG | 52 | 2233-2261 | 190 |
| M1-R | Reverse | TTA AAA AAA TTA AAT TAA TTC TTC C | 52 | 2423-2399 | |
| 2233-L | Forward | CAG TCT AGC AGT CAC CTT GGT GAT | 52 | 2235-2258 | |
| 2424-R | Reverse | CTG GAG AGT TGA GAA AGT TGG GTT | 52 | 2431-2408 | |
| M2-L | Forward | TTT AGT TAT TGT TTT TTA TGA TAT AGG AG | 54 | 3707-3735 | 298 |
| M2-R | Reverse | AAT TTT TCC ATC TAA TTC CAT CTA CTC | 54 | 4005-3979 | |
| 3694-L | Forward | GGG ATA AGC CTA TCT CAG TCA TT | 54 | 3964-3716 | |
| 3998-R | Reverse | CCA TCT GAT TCC ATC TGC TCC | 54 | 3998-3978 | |
| M3-L | Forward | GGT TAA TGG ATT AGA GAT GTA TTT AGG G | 59 | 6192-6219 | 241 |
| M3-R | Reverse | TTT ACT TTC CCT TTT CTT TTT CTT AAA A | 59 | 6433-6406 | |
| 6193-L | Forward | GGC CAA TGG ATT AGA GAT GCA CTT | 59 | 6192-6215 | |
| 6456-R | Reverse | CGC TTC TGG TGG ATG ATC TTG A | 59 | 6456-6434 | |
| M4-L | Forward | AGT GGT TTT ATG TGT TTA AAG TAT G | 50 | 6631-6655 | 339 |
| M4-R | Reverse | AAA AAT TAT TCA CAC AAA CAT TTT ATT A | 50 | 6970-6943 | |
| 38146(D) | Forward | AGT AGT GGT CCT ATG TGT C | 50 | 6628-6646 | 356 |
| 38148(D) | Reverse | GTT TGG AAA CAA GAG AGA GTTATTC | 50 | 6984-6960 | |
| For Chromatin immunoprecipitation and RT-PCR | | | | | |
| 35552 (A) | Forward | CAA GGT CTT ATC AGA GTT CTC G | 50 | 1773-1794 | 565 |
| 35555 (A) | Reverse | GAG GAT GTA TCT TCC GAA GG | 50 | 2338-2319 | |
| 34394 (B) | Forward | CGC ATT GGA GCA GAT GG | 50 | 3971-3987 | 482 |
| 34393 (B) | Reverse | GTG AGA GAA GAG TGT GAG | 50 | 4453-4436 | |
| 37230 (C) | Forward | TTC CAG CCC ATC CTT TTT TCA G | 50 | 4639-4660 | 683 |
| 34961 (C) | Reverse | TGC TCT CAT GTC CAT TTC AAC C | 50 | 5322-5301 | |
| 38146 (D) | Forward | GCT CTA GAA GTA GTG GTC CTA TGT GTC | 50 | 6628-6653 | 356 |
| 38148 (D) | Reverse | CAT GCC ATG GTT TGG AAA CAA GAG AGA GTT ATT C | 50 | 6984-6960 | |
| Rpl2(F) | Forward | GA GCG AAA TGG GTA CCT CAA G | 58 | | 400 |
| Rpl2(R) | Reverse | CCA GCA ACC TGA CCA ATC ATA G | 58 | | |
| Tph1-F | Forward | AAG GGG GTT CAA GAG GAC AAC CGT | 58 | | 250 |
| Tph1-R | Reverse | CTT GGT GCA AAG GGG TGT CAA TTG | 58 | | |
| For siRNA analysis | | | | | |
| 2233-L | Sense | CAG TCT AGC AGT CAC CTT GGT GAT | | 2235-2258 | |
| 3694-L | Sense | GGG ATA AGC CTA TCT CAG TCA TT | | 3694-3716 | |
| 2233-as | Anti-sense | ATC ACC AAG GTG ACT GCT AGA CTG | | 2258-2235 | |
| 3694-as | Anti-sense | AAT GAC TGA GAT AGG CTT ATC CC | | 3716-3694 | |
| 6798-s | Sense | ACT ATC GAG TCT GTA TTA TTA AGT GTG ATA | | 6798-6827 | |
| 6798-as | Anti-sense | TAT CAC ACT TAA TAA TAC AGA CTC GAT AGT | | 6827-6798 | |
| *dTph*1-s | Sense | TGG CAG CCC AGA TTC GAT CCT GGG TGG TGA | | | |
| *dTph1*-as | Anti-sense | TCA CCA CCC AGG ATC GAA TCT GGG CTG CCA | | | |
| *dTph1*-2s | Sense | ACA CAC AGT ATA TAT CCA ATC TAA ACC AAT | | | |
| *dTph1*-2as | Anti-sense | ATT GGT TTA GAT TGG ATA TAT ACT GTG TGT | | | |

**Table 1.** Sequences of oligonucleotides used, indicating their position within the PVCV genome if applicable.

viral DNA nor viral particles can be detected. In the symptomatic state, plants have symptoms of vein clearing, and episomal viral DNA and particles can be detected (Richert-Pöggeler et al., 2003). By using the sensitive detection method of RT-PCR, we were able to detect viral transcripts in both states.

Transcripts arising from circular episomal virus or from tandemly integrated copies would be full-length, i.e. they would include the viral promoter region, while truncated transcripts derived from the PVCV promoter would not (Fig. 1b). Accordingly, in non-symptomatic plant material we found transcripts corresponding to the three different coding regions tested (Fig. 2a; primer pairs A, B and C), while transcripts from the promoter region (primer pair D) were below detection levels. Even from the coding region the amount of transcripts was low, although levels increased with plant age (Fig. 2b). Symptomatic plants, however, gave rise to considerably higher levels of transcripts, including transcripts covering the intergenic region (Fig. 2a, primer pairs A-D).

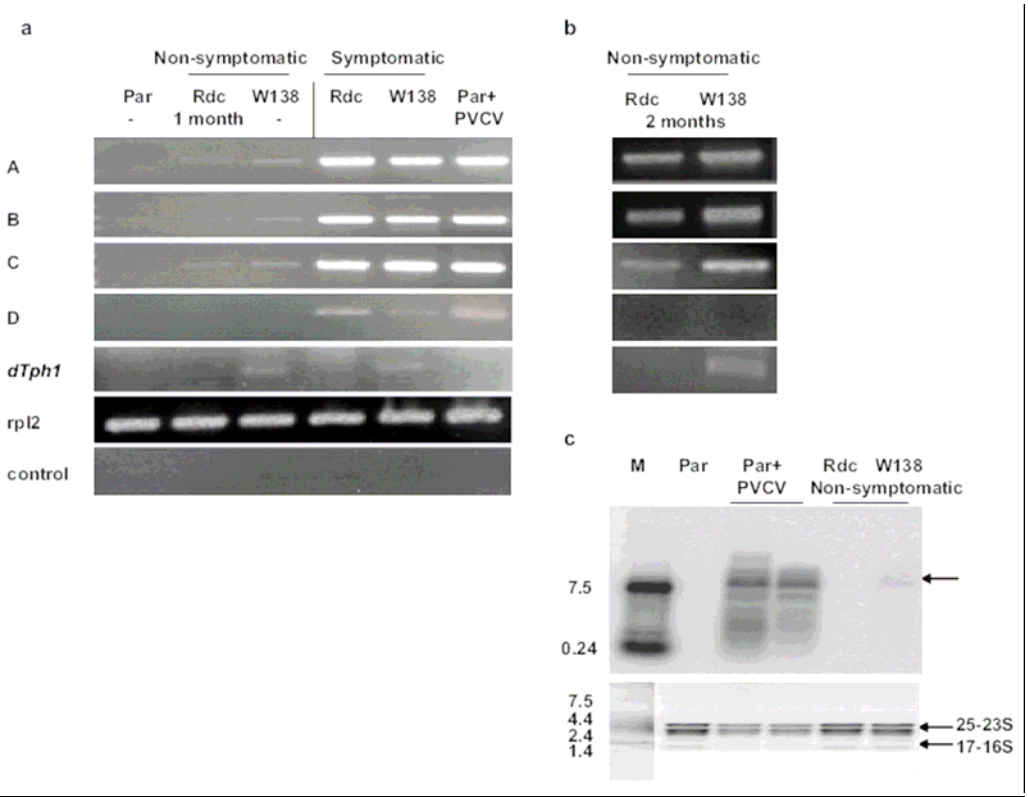

**Fig 2.** a) RT-PCR of *P. parodii* ePVCV-1 free (Par), 1-month-old non-symptomatic and symptomatic *P. hybrida* Rdc and W138 and infected *P. parodii* with the full-length PVCV infectious clone (Par+PVCV). Primer pairs A, B, C and D (Fig. 1) were used. Ribosomal protein large subunit 2 (rpl2) RNA served as an internal control. A *dTph1* transposon-specific product was also analyzed. The control panel shows the result of RT-PCR with the omission of reverse transcription using primer pair (rpl2). b) RT-PCR of non-symptomatic and symptomatic 2-month-old *P. hybrida* Rdc and W138 plants (primer pairs A-D). c) Northern analysis of RNA from PVCV-infected plants and plants harbouring ePVCV-1. Par; *P. parodii* (ePVCV-free), Par+PVCV; *P. parodii* infected with the full-length PVCV infectious clone, M; Size marker. The arrow indicates the cloud of transcripts in W138. The lower panel shows the ethidium bromide-stained gel, showing ribosomal RNA.

From these findings, we conclude that integrated viral genomes are silenced, but that this silencing is incomplete, allowing a low level of transcription. Transcripts carrying PVCV-1 promoter sequences would be expected to be derived from tandemly arranged ePVCV-1 copies. Since those were below detection limit, we assume that silencing of tandemly arranged copies is tighter than that of at least some of the defective integrants.

Interestingly, the ePVCV-1 sequences proved to be more transcriptionally active in the W138 than in the Rdc background (Fig. 2b). In addition, northern assays were sensitive enough to visualize clouds of transcripts in W138 but not in Rdc (Fig. 2c).

### 4.2: A bias against methylation of the ePVCV-1 intergenic region

The evident leakiness of transcriptional silencing in W138 correlates with the high efficiency of *dTph1* transposition in this cultivar. Indeed, a *dTph*1-specific RT-PCR product was detected only in W138. To investigate the underlying cause of this higher level of transposition in more detail, studies on ePVCV-1 and *dTph1* DNA methylation and accompanying histone modification were performed.

Using methylation-sensitive restriction enzymes, we had previously shown that ePVCV-1 sequences in Rdc are controlled by the host via cytosine-methylation (Richert-Pöggeler et al., 2003). To generate C-methylation maps of selected parts of the ePVCV-1 coding region to single base resolution, genomic DNA isolated from 1-month-old *P. hybrida* cultivars Rdc and W138 was used for bisulfite sequencing of representative domains of ~200 bp long using primer pairs

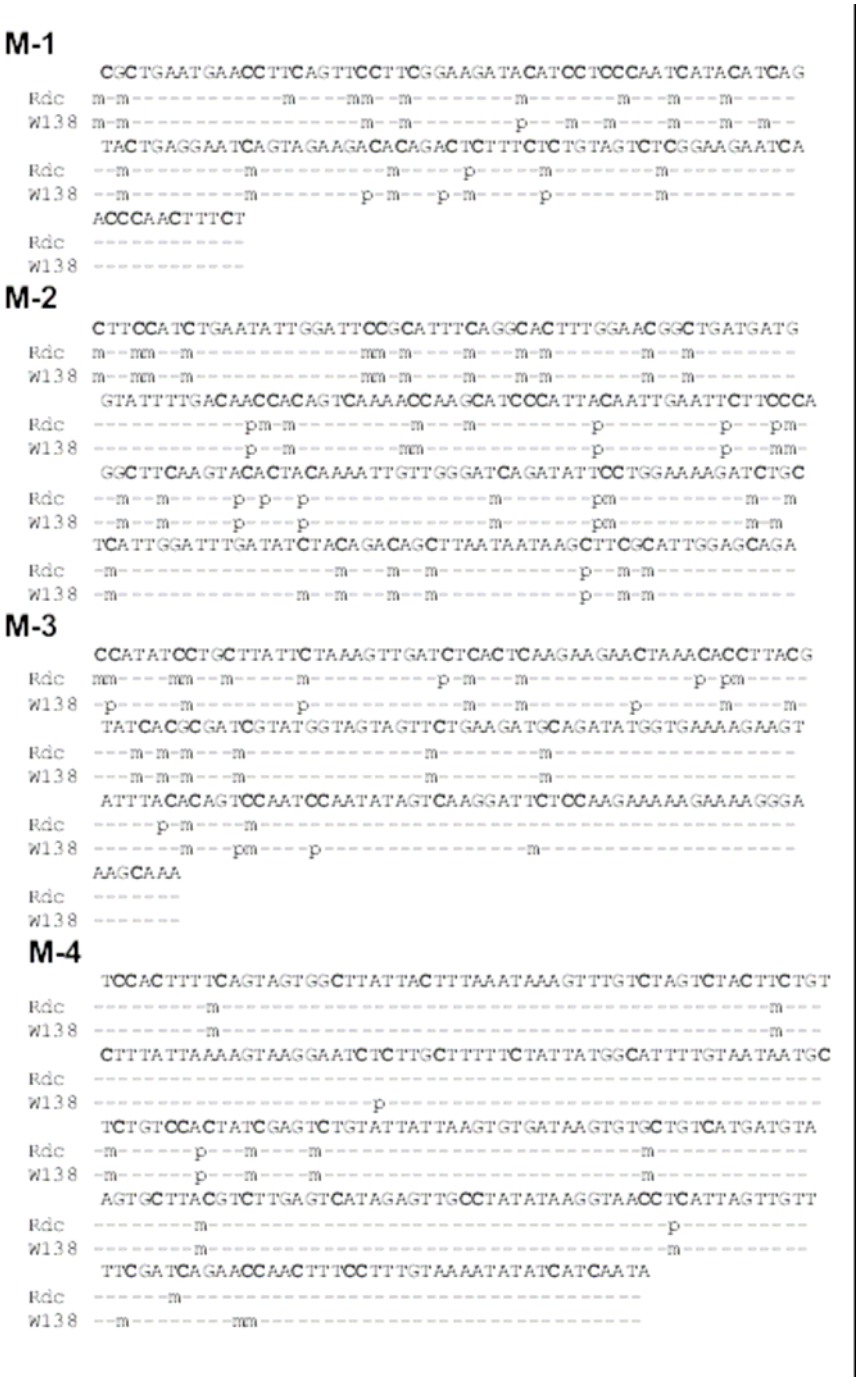


**Fig 3.** Sequence comparison of bisulfite-treated DNA from Rdc and W138, showing local differences. Cytosine residues (indicated as bold C) were classified as methylated (m), partially methylated (p), or unmethylated (-). When 90-100%, 50%, or 0-10% of the cytosines at a certain position were methylated.

M-1, -2 and -3 (Fig. 1). In both cultivars, more than 50% of the cytosine residues were found to be methylated (Fig. 3 and Table 2). Methylation was the predominant state in symmetrical context (CNG and CG), and less pronounced in the asymmetric context (CNN). No significant differences between Rdc and W138 were observed. Despite the overall similarity in methylation pattern of the two cultivars, some local differences in cytosine methylation were revealed in all sequence contexts (Fig. 3).

| **Analyzed region** | **M1** | | | **M2** | | | **M3** | | | **M1 to M3** | | | **M4** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Rdc** | **total** | **metC** | **%** | **total** | **metC** | **%** | **total** | **metC** | **%** | **total** | **metC** | **%** | **total** | **metC** | **%** |
| CG | 3 | 3 | 100 | 3 | 3 | 100 | 4 | 3 | 75 | 10 | 9 | 90 | 3 | 2 | 66 |
| CNG | 7 | 6 | 86 | 11 | 10 | 91 | 4 | 4 | 100 | 22 | 20 | 91 | 6 | 6 | 100 |
| CNN | 23 | 6 | 26 | 33 | 16 | 49 | 24 | 10 | 42 | 80 | 32 | 40 | 34 | 0 | 0 |
| Sum | 33 | 15 | 45 | 47 | 29 | 62 | 32 | 17 | 53 | 112 | 61 | 54 | 43 | 8 | 19 |
| | | | | | | | | | | | | | | | |
| W138 | | | | | | | | | | | | | | | |
| CG | 3 | 3 | 100 | 3 | 3 | 100 | 4 | 4 | 100 | 10 | 10 | 100 | 3 | 3 | 100 |
| CNG | 7 | 5 | 71 | 11 | 10 | 91 | 4 | 4 | 100 | 22 | 19 | 86 | 6 | 5 | 83 |
| CNN | 23 | 6 | 26 | 33 | 17 | 52 | 24 | 6 | 25 | 80 | 29 | 36 | 34 | 3 | 9 |
| Sum | 33 | 14 | 42 | 47 | 30 | 64 | 32 | 14 | 44 | 112 | 58 | 52 | 43 | 11 | 26 |

**Table 2.** Cytosine methylation pattern after bisulfite sequencing of selected regions of coding and non-coding regions of integrated PVCV sequences. Total numbers of cytosines (total) are compared with numbers of methylated cytosines (metC). Primer pairs M1, M2, and M3 cover the coding region, with M4 covering the non-coding region (see Fig. 1).

### 4.3: Histone 3 (H3) methylation of endogenous PVCV sequences

More pronounced differences between Rdc and W138 were observed by chromatin immunoprecipitation (ChIP). Chromatin was fixed with formaldehyde and histone-DNA complexes were precipitated using anti-H3 -dimethyl K4 (anti-H3mK4) and anti-H3 dimethyl K9 (anti-H3mK9)-specific antibodies, representative of active and inactive chromatin, respectively. PCR was performed with primers specific to the coding region of the ribosomal protein gene rpl2, representing active chromatin, for the transposon *dTph1*, usually representing inactive chromatin, and for the ePVCV-1 sequences to be analysed. ePVCV-1-specific primer annealing sites were located within the coding region (A, B, C in Fig. 1) or within the promoter region (D), including 70 bp of sequence downstream of the transcription start site (F. Noreen et al., unpublished). As expected, rpl2 DNA was predominantly associated with H3mK4 in both cultivars (Fig. 4). In contrast, *dTph1* was associated with H3mK9 in Rdc and with H3mK4

in W138, indicating a deficiency in transposon silencing in the latter cultivar. Analysis of the ePVCV-1 coding region revealed H3mK9 in Rdc and a mixture of H3mK9 and H3mK4 in W138, again indicating a silencing deficiency in the latter cultivar.

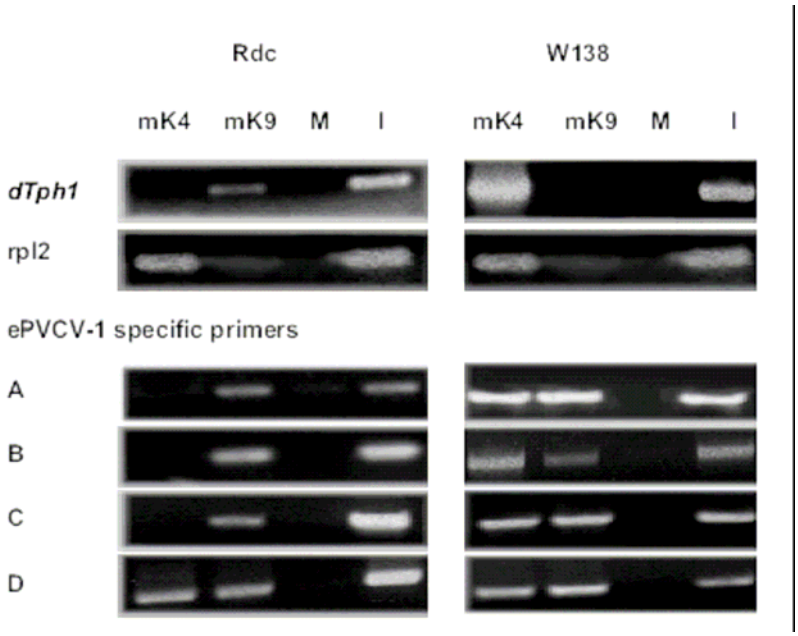


**Fig 4.** Chromatin immunoprecipitation (ChIP) of *P. hybrida* Rdc and W138. Primer sets specific for rpl2 (control for active chromatin), *dTph1*, (inactive in Rdc and active in W138) and ePVCV-1-specific primers from coding (A, B, C) and intergenic (D) regions were used to amplify the PCR products from total input (I), no-antibody mock precipitation control (M), chromatin immunoprecipitated with H3 anti-dimethyl K4 antibodies (mK4), or chromatin immunoprecipitated with H3 anti-dimethyl K9 antibodies (mK9) from the *P. hybrida* cultivars indicated.

Analysis of the ePVCV-1 promoter region revealed striking differences with the coding region: asymmetric cytosines were less methylated (Fig. 3; primer pair M4) and histones had a higher proportion of H3mK4 (Fig. 4). This indicates that a certain level of transcription is required to initiate *de novo* DNA methylation and histone modification of ePVCV-1 coding region as visible in Rdc (Fig. 2a and Fig. 4). It also suggests that the PVCV promoter region has evolved to avoid being completely silenced by histone modification. Interestingly, all ePVCV-1 sequences analysed in W138 were associated equally with H3mK9 and H3mK4, implying a less efficient heterochromatization potential in this petunia line. This is even more obvious for the DNA transposon *dTph1*, which is found exclusively in the H3mK4 fraction.

### 4.4: siRNAs

siRNAs are the hallmarks of silencing. In *Arabidopsis thaliana* three size classes have been found, indicative of the actions of several individual dicer-like (DCL) genes (Park et al., 2002; Xie et al., 2004; Vazquez et al., 2004; Dunoyer et al., 2005; Gasciolli et al., 2005; Akbergenov et al., 2006). DCL3 produces 24nt-long siRNAs thought to be involved in TGS; DCL4, 21nt-long bands involved in PTGS; DCL2, 22nt-long bands of as yet unclear function); and DCL1, 21nt or longer bands involved in microRNA (miRNA) production. To study the involvement of siRNAs in ePVCV-1 silencing, their presence was analyzed in northern hybridizations using oligonucleotides specific to the virus genome coding and intergenic region in the sense and antisense orientation as probes. No siRNAs were detected for the ePVCV-1-free line *P. parodii* (Fig. 5) used as a negative control. In symptomatic plants generated by inoculation with cloned viral DNA or by induction from endogenous ePVCV-1 copies, siRNA pools of all three size classes (21, 22 and 24nt) originating both from the coding and intergenic region were readily detectable (Fig. 5). Both sense and antisense probes hybridized to the membrane with equal strength, implying the existence of sense and antisense siRNAs, and dsRNA molecules as precursors. For siRNAs produced from the intergenic region, the sense probe generally gave rise to a stronger signal than the antisense probe. We could not detect any siRNA derived from the intergenic region in non-symptomatic plants even when the RNA samples were enriched for smaller sized RNA molecules. Assuming that the types of DCL-enzymes and their products are similar in *A. thaliana* and petunia, and ignoring miRNAs, this indicates that both TGS, marked by the 24nt species, and PTGS, marked by the 21 and 22nt species, occur. siRNA levels in non-symptomatic tissue were very low or below the limit of detection. Thus, siRNA-levels are an excellent indicator of the state of induction of ePVCV-1.

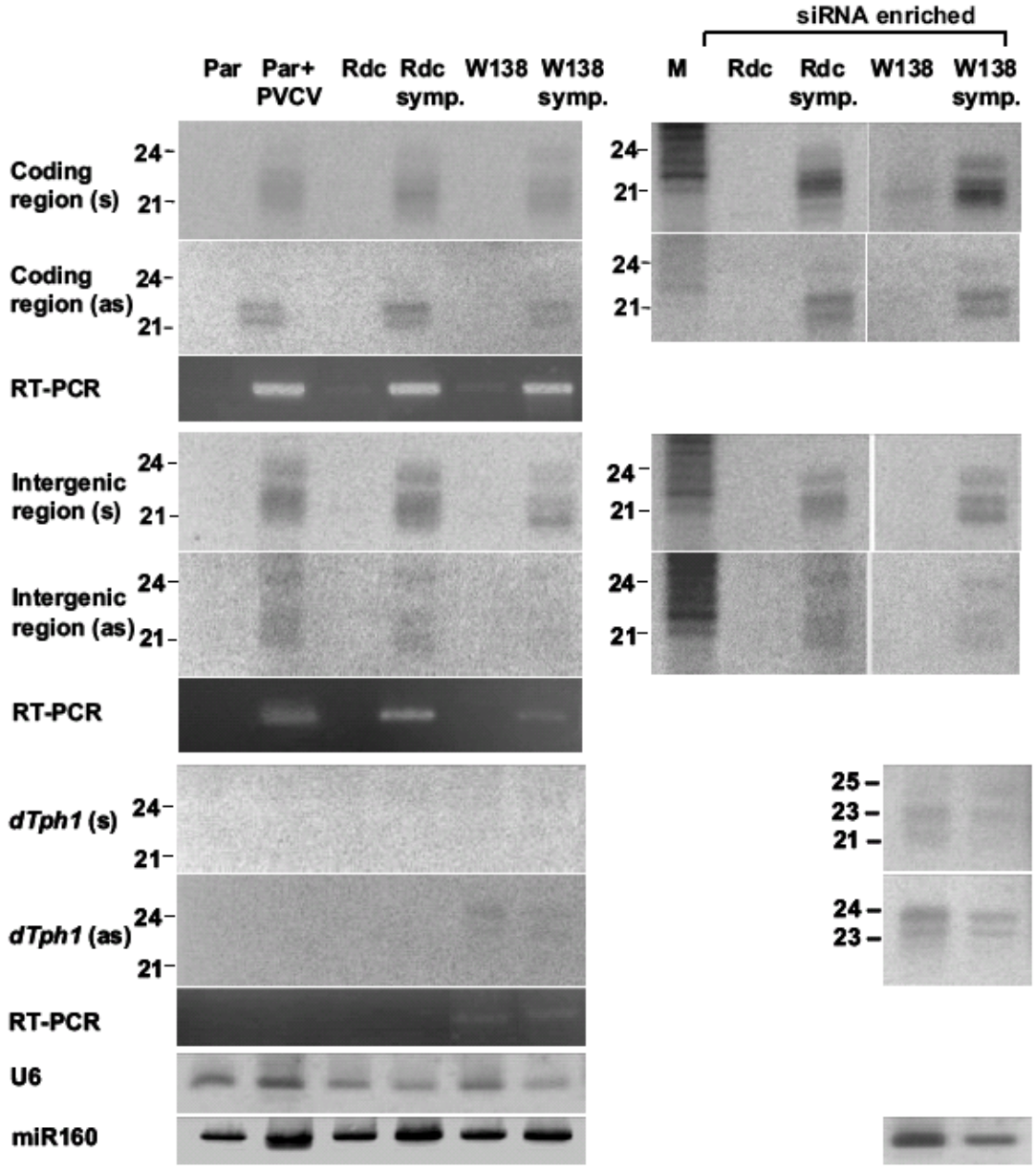


**Fig 5:** Northern hybridisation for siRNA detection in symptomatic and non-symptomatic petunia. The membranes of the right and left panel were probed and rehybridised several times with different oligonucleotides. Marker (M). Sizes of 21 and 24 nt RNAs are indicated. *P. parodii* ePVCV-1-free (Par) and infected *P. parodii* (Par+PVCV) were used as negative and positive controls for the presence of episomal virus. Oligonucleotides of sense (s) and antisense (as) orientation were used as probes for the coding and intergenic regions. “siRNA enriched” depicts plant material that has been additionally enriched for small RNAs after total RNA isolation. Detection of transposon (*dTph1*)-specific siRNAs using sense and antisense oligonucleotides as probes (see Table 1). For better comparison corresponding RT-PCR results of Fig. 2 (RT-PCR) are inserted below the siRNA. As control for equal loading the blots were probed with U6-RNA- and miR160-specific oligonucleotides in the case of total RNA preparations and with the miRNA160 specific probe in the case of enriched RNA preparations.

As expected from the RT-PCR and ChIP assays, transposon (*dTph1*)-specific siRNAs of 21-25 nt in size and in both orientation were produced only in W138, which exhibits active transcription of this element. The amount of siRNAs detected was indifferent to viral infection.

## 5: Discussion

The major part of eukaryotic chromatin exists in a condensed and transcriptionally silenced state. In plants and many other organisms, heterochromatic DNA is characterized by cytosine methylation in symmetric and asymmetric context, and by specifically modified histones (Almeida and Allshire, 2005), best exemplified by histone 3 lysine 9 dimethylation (H3mK9) and lack of histone 3 lysine 4 methylation (H3mK4). Large portions of such heterochromatin are located in the pericentromeric regions of the chromosomes, which are also the preferred locations of mobile elements and of defective and intact endogenous pararetroviruses, such as ePVCV-1 of *Petunia hybrida* (Richert-Pöggeler et al., 2003). In these positions, parasitic elements cause only minimal harm to the host, while they themselves survive and are passively replicated.

The complex pathway required for transcriptional gene silencing (Tariq and Paszkowski, 2004; Wassenegger, 2005) is thought to include residual transcription directed by POL II and/or POL IV (Herr et al., 2005); production of dsRNA by RNA-dependent RNA polymerase RDR2 (Sugiyama et al., 2005) and of 24nt cognate siRNA by DCL3 (Gasciolli et al., 2005); formation of an RNAi-induced transcriptional silencing (RITS) complex (Verdel et al., 2004); RNA-dependent DNA methylation (RdDM; Aufsatz et al., 2002; Kanno et al., 2005) involving *de novo* methylation of potentially all cytosines (Cao et al., 2003; Chan et al., 2004; Matzke and Birchler, 2005) and maintenance methylation of those in symmetrical context (Takeda and Paszkowski, 2005); chromatin remodeling (Amadeo et al., 2000; Tariq et al., 2002) and chromatin modification, best exemplified by H3mK9 methylation directed by “KRYPTONITE” (KYP) (Jackson et al., 2002) and lack of H3mK4 methylation (Lee et al., 2005).

Our analysis reveals that TGS of ePVCV-1 conforms to this path, i.e. we observed residual transcription, siRNA formation, DNA methylation at

asymmetric and symmetric cytosines, and appropriate histone modification. However, we also observed three types of partial deviation from this scheme: 1) coding and promoter regions of ePVCV-1 are differentially affected; 2) the two petunia cultivars, Rdc and W138, differ in the degree of histone modification of the coding region; and 3) there is a drastic increase in siRNA concentration upon induction of the provirus and establishment of episomal PVCV-1.

The ePVCV-1 coding region was characterized as having a significant amount of transcripts, a high level of overall cytosine methylation, and histone H3mK9 methylation. In contrast, transcripts from the promoter region were below our detection limit, the level of cytosine methylation was relatively low, and the accompanying histones have a lower H3mK9/H3mK4 ratio than those from the coding region. One reason might be the different availability of ePVCV-1 coding and promoter regions for transcription due to the distribution and types of integrants as discussed in the Results section. If the promoters of single copy inserts are used for low-level transcription, then the promoters themselves are not transcribed and only minor amounts of 24nt siRNAs would be available to direct cognate chromatin modification. In line with this hypothesis, we found that the e-PVCV-1 promoter regions are less methylated, and are associated with a higher proportion of H3mK4 than the coding regions. This effect would be markedly different for integrated retrotransposons, where promoters are located in the long terminal repeats (LTRs) of the integrants, with the upstream LTR promoting transcription and the downstream one being transcribed.

The H3mK9/H3mK4 ratio, but not the level of DNA methylation, of e-PVCV-1 sequences was markedly lower in petunia line W138 than in Rdc, despite the comparable copy number in the two cultivars (K. Richert-Pöggeler and T. Schwarzacher, unpublished). Since transcript- and siRNA-levels were similar for the two lines, we assume that a deficiency in some late state of the TGS pathway, such as siRNA usage, chromatin remodeling, KYP activity and/or an impaired H3mK4 demethylase (Lee et al., 2005) is responsible for these effects. Such a deficiency might also be responsible for the high activity of

transposon *dTph1* in W138 and the considerable amounts of the corresponding transcripts and siRNAs detected in our assays.

Once the methylation pattern is disturbed, e.g. upon induction (Richert-Pöggeler et al., 2003), pregenomic RNA with terminal repeats including promoter sequences can be produced. Upon reverse transcription, the resulting DNA accumulates as episomal circular DNA. As a result a significant increase in transcripts as well as siRNA production compared to the uninduced case can be observed. This increase also leads to a drastic increase in siRNA levels, although apparently this is not sufficient to block virus amplification.

What makes the ePVCV-1 provirus state relatively stable, giving rise to infectious outbursts only rarely and under stress conditions? One reason might be that most ePVCV-1 transcripts are defective and have to await repair by corrective mutation and/or recombination. Another possibility might be that full infection depends on transcript thresholds; we suggest that a certain balance between silencing and silencing suppression exists. The presence of many transcripts that are incomplete because they originate from random integration events interrupting the viral genome, and because of their lack of terminal repeats, might push the equilibrium towards the silenced state. Nevertheless, an increasing number of complete transcripts will accumulate over time, derived either from recombinational repair of mutants or transcription from intact tandem integrants. Once a certain threshold of such intact transcripts is reached, sufficient amounts of suppressor and movement protein are produced to counteract silencing and to allow viral spread to other cells, respectively. One has to keep in mind that the thresholds necessary for infection from within are not reached synchronously in each cell, and symptom expression will be observed only if episomal virus can be established in a number of neighbouring cells. This would explain the long latent period of infection and the stochastic expression of symptoms. Abiotic or biotic factors that lead to genomic stress and disturbance of host defence systems could be advantageous for the virus to establish infection. Indeed, PVCV symptoms have been found more frequently when plants were

kept under higher temperatures or under water stress (Lockhart and Lesemann, 1997; Zeidan et al., 2001).

In conclusion, we have shown that in *P. hybrida*, as has been reported for *Arabidopsis thaliana*, different sized siRNAs are produced, thus indirectly providing evidence for the presence of multiple dicers involved in TGS and PTGS (DCL1-4; Xie et al., 2004). Such cooperative efforts of different RNAi pathways in the control of foreign nucleic acids are not surprising, since the life cycle of ePVCV-1 as well as PVCV-1 includes nuclear as well as cytoplasmic stages.

## 6: References


**Akbergenov, R., Si-Ammour, A., Blevins, T., Amin, I., Kutter, C., Vanderschuren, H., Zhang, P., Gruissem, W., Meins, F., Hohn, T., and Pooggin, M.M.** (2006). Molecular characterization of geminivirus-derived small RNAs in different plant species. Nucleic Acids Res. **in press**.

**Almeida, R., and Allshire, R.C.** (2005). RNA silencing and genome regulation. Trends Cell Biol **15,** 251-258.

**Amedeo, P., Habu, Y., Afsar, K., Mittelsten Scheid, O., and Paszkowski, J.** (2000). Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature **405,** 203-206.

**Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, A.J., and Matzke, M.** (2002). RNA-directed DNA methylation in Arabidopsis. Proc Natl Acad Sci U S A **99 Suppl 4,** 16499-16506.

**Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E.** (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol **13,** 2212-2217.

**Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., and Jacobsen, S.E.** (2004). RNA silencing genes control de novo DNA methylation. Science **303,** 1336.

**Dunoyer, P., Himber, C., and Voinnet, O.** (2005). DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA

component of the plant cell-to-cell silencing signal. Nat Genet **37,** 1356-1360.

**Gasciolli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H.** (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol **15,** 1494-1500.

**Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A.** (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science **297,** 1871-1873.

**Gerats, A.G., Huits, H., Vrijlandt, E., Marana, C., Souer, E., and Beld, M.** (1990). Molecular characterization of a nonautonomous transposable element (dTph1) of petunia. Plant Cell **2,** 1121-1128.

**Gerats, T., and Vandenbussche, M.** (2005). A model system for comparative research: Petunia. Trends Plant Sci **10,** 251-256.

**Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G.** (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature **386,** 44-51.

**Grandbastien, M.A.** (2004). [Stress activation and genomic impact of plant retrotransposons]. J Soc Biol **198,** 425-432.

**Harper, G., Osuji, J.O., Heslop-Harrison, J.S., and Hull, R.** (1999). Integration of banana streak badnavirus into the Musa genome: molecular and cytogenetic evidence. Virology **255,** 207-213.

**Harper, G., Hull, R., Lockhart, B., and Olszewski, N.** (2002). Viral sequences integrated into plant genomes. Annu Rev Phytopathol **40,** 119-136.

**Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C.** (2005). RNA polymerase IV directs silencing of endogenous DNA. Science **308,** 118-120.

**Hirochika, H., Okamoto, H., and Kakutani, T.** (2000). Silencing of retrotransposons in arabidopsis and reactivation by the ddm1 mutation. Plant Cell **12,** 357-369.

**Hohn, T., and Richert-Pöggeler, K.R.** (2005). Replication of plant pararetroviruses. In: Hefferon, K.R. (ed) Recent advances in DNA virus replication. Research Singpost Transworld (Kerala, India).

**Hughes, J.F., and Coffin, J.M.** (2001). Evidence for genomic rearrangements mediated by human endogenous retroviruses during primate evolution. Nat Genet **29,** 487-489.

**Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E.** (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416,** 556-560.

**Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E.** (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma **112,** 308-315.

**Johnson, L., Cao, X., and Jacobsen, S.** (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Curr Biol **12,** 1360-1367.

**Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J.** (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet **37,** 761-765.

**Kunii, M., Kanda, M., Nagano, H., Uyeda, I., Kishima, Y., and Sano, Y.** (2004). Reconstruction of putative DNA virus from endogenous rice tungro bacilliform virus-like sequences in the rice genome: implications for integration and evolution. BMC Genomics **5,** 80.

**Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R.** (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature **437,** 432-435.

**Lesemann, D.-E., and Casper, R.** (1973). Electron microscopy of petunia vein clearing virus, an isometric plant virus associated with specific inclusions in petunia cells. Phytopathology **63,** 1118-1124.

**Li, L.C., and Dahiya, R.** (2002). MethPrimer: designing primers for methylation PCRs. Bioinformatics **18,** 1427-1431.

**Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh, S., and Jacobsen, S.E.** (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. Embo J **23,** 4286-4296.

**Lockhart, B.E., Menke, J., Dahal, G., and Olszewski, N.E.** (2000). Characterization and genomic analysis of tobacco vein clearing virus, a plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. J Gen Virol **81,** 1579-1585.

**Lockhart, B.E.L., and Lesemann, D.-E.** (1997). Occurrence of petunia vein clearing virus in the U. S. A. Plant Dis **82,** 262.

**Mathieu, O., Probst, A.V., and Paszkowski, J.** (2005). Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in Arabidopsis. Embo J **24,** 2783-2791.

**Matzke, M.A., and Birchler, J.A.** (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet **6,** 24-35.

**Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T.** (2001). Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature **411,** 212-214.

**Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olszewski, N.E., and Lockhart, B.** (1999). Evidence that badnavirus infection in Musa can originate from integrated pararetroviral sequences. Virology **255,** 214-220.

**Park, W., Li, J., Song, R., Messing, J., and Chen, X.** (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol **12,** 1484-1495.

**Richards, E.J.** (1997). DNA methylation and plant development. Trends Genet **13,** 319-323.

**Richert-Pöggeler, K.R., Noreen, F., Schwarzacher, T., Harper, G., and Hohn, T.** (2003). Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. Embo J **22,** 4836-4845.

**Stuurman, J., and Kuhlemeier, C.** (2005). Stable two-element control of dTph1 transposition in mutator strains of Petunia by an inactive ACT1 introgression from a wild species. Plant J **41,** 945-955.

**Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I.** (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A **102,** 152-157.

**Takeda, S., and Paszkowski, J.** (2005). DNA methylation and epigenetic inheritance during plant gametogenesis. Chromosoma**,** 1-9.

**Tanne, E., and Sela, I.** (2005). Occurrence of a DNA sequence of a non-retro RNA virus in a host plant genome and its expression: evidence for recombination between viral and host RNAs. Virology **332,** 614-622.

**Tariq, M., and Paszkowski, J.** (2004). DNA and histone methylation in plants. Trends Genet **20,** 244-251.

**Tariq, M., Habu, Y., and Paszkowski, J.** (2002). Depletion of MOM1 in non-dividing cells of Arabidopsis plants releases transcriptional gene silencing. EMBO Rep **3,** 951-955.

**Van Houwelingen, A., Souer, E., Spelt, K., Kloos, D., Mol, J., and Koes, R.** (1998). Analysis of flower pigmentation mutants generated by random transposon mutagenesis in Petunia hybrida. Plant J **13,** 39-50.

**Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P.** (2004). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Mol Cell **16,** 69-79.

**Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D.** (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science **303,** 672-676.

**Walbot, V.** (1999). UV-B damage amplified by transposons in maize. Nature **397,** 398-399.

**Wassenegger, M.** (2005). The role of the RNAi machinery in heterochromatin formation. Cell **122,** 13-16.

**Xie, Z., Allen, E., Wilken, A., and Carrington, J.C.** (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. Proc Natl Acad Sci U S A **102,** 12984-12989.

**Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C.** (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol **2,** E104.

**Zeidan, M., Sikron, N., Cohen, J., and Gera, A.** (2001). Improved detection of petunia vein clearing caulimovirus. HortScience **35**.

# Chapter 3

## Isolation and characterization of the promoter from an inducible endogenous plant pararetrovirus in petunia, *Petunia vein clearing virus* (PVCV)

Faiza Noreen[1], Suchandra Deb Roy[2], Matthias Müller, Neera Bhalla Sarin[2], Thomas Hohn[1], Katja Richert-Pöggeler[1].
[1]University of Basel, Botanical Institute, Plant Health Unit, schönbeinstrasse 6, 4056 Basel, Switzerland. [2]School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

(Manuscript in preparation)

## 1: Summary

Promoter- and polyadenylation-sequences of *Petunia vein clearing virus* (PVCV), a pararetrovirus present in endogenous and episomal forms, were mapped. The promoter, including upstream and downstream enhancer elements, was cloned and used to drive the expression of a *β-glucuronidase* (GUS) reporter gene in plant protoplasts and tobacco plants, showing activity in root, leaf and stem. Although originally of medium strength when compared to the widely used 35S promoter of *cauliflower mosaic virus* (CaMV), the promoter could be improved to about 50% the CaMV 35S promoter strength by reconstitution of repetitive sequences within the as-1 like enhancer element of the PVCV promoter. This novel promoter may be highly valuable for plant transformation system requiring heterologous promoters for gene expression to avoid transcriptional gene silencing based on homologous sequence recognition.

## 2: Introduction

Plant pararetrovirus genomic promoters interact with plant RNA polymerase II (Olszewski et al., 1982) and their regulation is apparently not dependent on any virus-encoded transcriptional factors (Odell et al., 1985). The strength of these promoters might have evolved to guarantee abundant production of terminally redundant pregenomic RNA (pgRNA) which serves as both, polycistronic mRNA and as template for replication via reverse transcription, thus rate limiting the viral replication (Rothnie et al., 1994). Consequently, these promoters are capable of causing high levels of gene expression in transgenic plants. The 35S promoter from CaMV has been widely used in chimeric gene constructs and is well characterized (Benfey et al., 1990). Matching promoters from *figwort mosaic virus* (FMV), *peanut chlorotic streak virus* (PCISV) and *cestrum yellow leaf curling virus* (CmYLCV) are also constitutively active in transgenic plants (Sanger et al., 1990; Stavolone et al., 2003; Hohn and Richert-Poggeler, 2005, submitted). Promoters isolated from bacilliform DNA viruses, unlike caulimovirus (isometric particles) promoters, seems to be primilary active in vascular tissues. Indeed it has been reported that c*ommelina yellow mottle virus* (CoYMV) and *rice tungro bacilliform virus* (RTBV) promoters direct phloem-specific gene expression in transgenic plants (Medberry et al., 1992; Bhattacharyya-Pakrasi et al., 1993). Promoters for pregenomic RNA of *banana streak virus* (BSV) are active for transgene expression in monocot and dicot plants (Schenk et al., 2001).

The regulatory region of the CaMV 35S promoter is made up of distinct domains that are involved in organ specific expression, each, but *in summa* cause the constituitivity by additive and synergistic interactions between these different domains (Benfey et al., 1990). Within these promoter domains several cis-acting elements were characterized. The activating sequence 1 (as-1) of CaMV that shows high homology to the palindrome of the ocs-element within T-DNA of *Agrobacterium tumefaciens* (Bouchez et al., 1989), incorporates a

TGACG repeat motif that is the binding site for the tobacco transcriptional factor TGA1a (Katagiri et al., 1989). It has been shown that this element has an important role in the control of gene expression in roots (Lam et al., 1989) and functions in synergy with upstream domains to potentiate promoter activities in other tissues (Fromm et al., 1989; Benfey et al., 1990). Plant promoters carrying as-1 like sequences have been reported to be involved in pathogen defense pathways and to react on induction by auxin and salicylic acid (Niggeweg et al., 2000). The as-1-binding factor (ASF-1) is composed mainly of the bZIP transcription factor TGA2.2 (80%) and to a lesser extend (10%) of TGA2.1 (Krawczyk et al., 2002). Whereas TGA2.2 and TGA2.1 activity is related to leaf tissue, another bZIP transcription factor TGA1a confers transcriptional activation in roots (Krawczyk et al., 2002). RTBV and CaMV enhancer elements were also found within the leader region, i.e. downstream of the transcription initiation site (Kloti et al., 1999; Pauli et al., 2004).

So far promoter activities from genera *caulimovirus* (CaMV, FMV, *mirabilis mosaic virus*; MMV and *blueberry red ringspot virus*; BRRV), *soymovirus* (*soybean chlorotic mosaic virus*; SbCMV, PCISV and CmYLCV), *cavemovirus* (*Cassava vein mosaic virus*; CsVMV), *badnavirus* (CoYMV, BSV, *sugarcane bacilliform badnavirus*; SCBV and *taro bacilliform virus*; TaBV) and tungroviruses (RTBV) have been tested (Glasheen et al., 2002; Yang et al., 2003). The present study was undertaken to identify the promoter region of PVCV, the type and (so far) only member of the family caulimoviridae genus *Petuviruses*. PVCV is distinct from other caulimoviruses because all its genetic information is comprised in one large open reading frame that suggests gene expression via a monocistronic RNA. We determined its ability to promote gene expression in transient assay using protoplasts and its tissue specificity using transgenic plants. Furthermore, we analyze the regulatory elements in this particular virus in comparison with those of other promoters isolated from plant pararetroviruses particularly functionality of as-1 like element (Richert-Poggeler and Shepherd, 1997) and its contribution in PVCV promoter expression. We show, that the promoter includes upstream and downstream enhancer elements, and that it can

be improved considerably by restoring the degenerated repetitive sequence of the as-1 like element. Transient expression assays did not show any silencing effect of homologous sequences present in the genome of *P. hybrida* on promoter activity in the episomal construct.

# 3: Materials and Methods

## 3.1: Plasmid construction

To identify the PVCV promoter fragments ranging from 321 to 1158bp were produced using the full-length PVCV clone p72-2 (Richert-Poggeler et al., 2003) as template for PCR that comprised deletions of upstream and downstream sequences with respect to the identified transcription start site. Three different reverse primers having *Nco*I site added at their 5' end and referred to as:

38147 "CATGCCATGGATAACTAGATCAGATGTTTTG" (PVCV genome position; 7146-7126),

38148 "CATGCCATGGTTTGGAAACAAGAGAGAGTTATTC" (PVCV genome position; 6984-6960),

38149 "CATGCCATGGTTTATTGATGATAATATTTTAC" (PVCV genome position; 6949-6927) were used. The corresponding forward primers contained an additional *Xba*I site at their 5' ends for cloning purposes and referred to as:

38146 " GCTCTAGAAGTAGTGGTCCTATGTGTC" (PVCV genome position; 6628-6653),

38145 " GCTCTAGAAGATCATCCACCAGAAGCGA" (PVCV genome position 6438-6457), and

38144 "GCTCTAGAAGACCAGTAAATTACTACCA (PVCV genome position; 5988-6007) were used.

| primer combination | produced fragment | fragment size(bp) | deletion |
|---|---|---|---|
| 38144/38147 | FNaa | 1158 | none |
| 38145/38147 | FNba | 708 | upstream |
| 38146/38147 | FNca | 518 | upstream |
| 38144/38148 | FNab | 996 | downstream |
| 38145/38148 | FNbb | 546 | up-, downstream |
| 38146/38148 | FNcb | 356 | up-, downstream |
| 38144/38149 | FNac | 961 | downstream |

| | | | |
|---|---|---|---|
| 38145/38149 | FNbc | 511 | up-, downstream |
| 38146/38149 | FNcc | 321 | up-, downstream |

**Table 1:** PVCV fragments isolated from non-coding region and tested for promoter activity.

PCR conditions were denaturation at 94 °C for 45 sec, annealing at 50 °C for 45 sec and extension at 72°C for 1 min with 30 cycles. The obtained fragments were cloned by replacing the *Xba*l-*Nco*l fragment covering the CaMV 35S promoter in plasmid p35S'-GUS (Hohn et al., 1996). In p35S'-GUS the *β-glucuronidase* (GUS) reporter gene is cloned into pUC19 vector (Stratagene) under the control of the CaMV 35S promoter and is terminated by CaMV poly (A) signal. This plasmid was used as reference for CaMV promoter activity. As an internal control a 35S-driven *chloramphenicol acetyltransferase (*CAT) gene construct was used. FNcb* was produced by introducing two mutations in as-1 like element of PVCV (TGATG$(n)_7$TTACG) restoring the palindrome of the as-1 element in CaMV (TGACG$(n)_7$TGACG) using PCR. FNcb plasmid was used ad the DNA template in this case. For DNA sythesis the forward primer "38146" and the reverse primer Rmas1 that include the mutations and a *BstE*II site for subcloning (PVCV genome position; 6896-6836, "TGAGGTTACCTTATATAGGC AATCTCTATGACTCAAGACGTCAGCACTT ACGTCATGAC") was used.

### 3.2: Isolation of protoplasts

Protoplast preparation of *Nicotiana plumbaginifolia* was performed as described previously by (Goodall et al., 1990). Petunia leaves (1 g) from plants grown in tissue culture were harvested and cut into small stripes (1 mm) in enzyme solution (see below). Tissues from midribs were discarded. For digestion of the cell wall the material was incubated overnight at 27 $^{o}$C in dark in 4 ml of sterile enzyme solution consisting of (for 100 ml): 1 g driselase (Fluka), 0.5 M mannitol, 0.25 g polyvinyl pyrrolidine, 0.1 g $CaCl_2$ x 2 $H_2O$, adjusted to pH 5.7 with KOH. After release of protoplasts during 10 min of gentle agitation and additional rest of 10 min, the preparation was filtered through a 100 µm pore size sieve and the sieve washed with an additional 10 ml of sterile solution W5 (per liter: 9 g NaCl, 18.3 g $CaCl_2$ x $2H_2O$, 0.37 g KCL, 0.99 g glucose adjusted to pH 6.0 with KOH). After centrifugation for 5 min at 800 rpm, the pellet was

resuspended in 10 ml of 0.6 M saccharose solution and overlayed with 1 ml of W5 solution and centrifuged (5 min at 800 rpm) was repeated and finally an aliquot taken for counting. Protoplast concentration was adjusted to 2x $10^6$/ml for transfection.

### 3.3: Transient expression assay

Polyethylene glycol (PEG)-mediated transfection of *N. plumbaginifolia* and petunia (*P. perodii*, *P. hybrida* lines Rdc and W138) were performed as described previously (Goodall et al., 1990). For all type of protoplasts 5 µg of test plasmid were routinely co-transfected together with 2 µg of a distinct plasmid used as an internal standard of transfection efficiency. GUS and CAT expression levels were determined upon protoplast cultivation and harvesting (Pooggin et al., 2000). Activities cited are an average of three independent transformations. CAT expression levels were determined using the CAT enzyme-linked immunosorbent assay kit (Roche Molecular Biochemicals) as recommended by the manufacturer. GUS activity was measured by a fluorimetric assay (Jefferson et al., 1986).

### 3.4: RNA isolation

For RNA isolation P. Parodii plants infected with full-length PVCV clone (p72-2) were used. Total RNA was extracted from infected leaves with the RNeasy Mini Kit (QIAGEN) with DNase 1 digestion of RNA according to the manufacturer's instructions.

### 3.5: 5’ and 3’ Rapid amplification of cDNA ends (RACE) analysis of RNA

The 5′ RACE system for Rapid Amplification of cDNA Ends (GIBCO BRL, Life Technologies, USA) was used for 5′ cDNA synthesis. As gene specific primer fn5 (5′- ATGGCTCTGATACCAATA-3′) was synthesized for first round of amplification and fn4 (5′- ATAACTAGATCAGATGTTTT-3′) was designed as nested primer for 5′ RACE. The PCR products were analyzed by sequencing.

cDNA synthesis was performed with the 3′ RACE system for Rapid Amplification of cDNA ends (RACE PCR Kit, GIBCO BRL, USA). An aliquot of 4 µg RNA was reversibly transcribed using primer AP (5′- GGCCACGCGTCGACTAG TAC$(T)_{17}$−3′) provided by the kit. Primer fn11 (5′-

GCTCTAGAAGACCAGTAA ATTACTACCA-3′) was designed according to the conserved amino acid sequence. Nested amplification was performed using primer fn10 (5′-GCTCTAGAAGATCATCCACCAGAAGCGA-3′). Reverse transcription (RT)-PCR was conducted according to the provided protocol. The PCR reaction was performed under the following conditions: cDNA was denatured at 94 °C for 3 min followed by 35 cycles of amplification (94 °C for 45 s, 57 °C for 1 min, 72 °C for 1 min) and by 72 °C for 10 min.

### 3.6: Cloning in binary vector and transformation of *Agrobacterium tumefaciens*

The *Xba*I-*SnaB*I fragment of plasmid FNcb and FNcb* was subcloned into pCAMBIA2301 (pCFNcb and pCFNcb*) placing the PVCV promoter sequences upstream the GUS reporter gene. The activity of both constructs was confirmed in transient transfection of protoplast (data not shown).

The binary vector was transferred into *A. tumefaciens* by freeze-thaw method. For the preparation of competent cells, *A. tumefaciens* (GV3101) was grown in 50 ml YEB (0.1% yeast extract, 0.5% beef extract, 0.5% peptone, 0.5% sucrose, 0.048 % $MgSO_4$) at 28 °C with vigorous shaking till the $O.D_{600}$ reached 0.6.Cells were pelleted at 3,500 rpm at 4 °C for 10 min. The pellet was resuspended in 50 ml 20 mM $CaCl_2$ and incubated in ice for 10 min and again pelleted down. This step was repeated with 25 ml of 50 mM $CaCl_2$. Finally, the pellet was resuspended in 10ml of 50 mM $CaCl_2$, aliquoted and stored at –80 $^oC$.

Approximately 1 µg of DNA was used for *A. tumefaciens* transformation. Competent cells were incubated with DNA for 30 min on ice and later dipped into liquid nitrogen. Heat shock was given at 37 $^oC$ for 5 min, followed by 2 min incubation on ice. This competent cell mix was incubated at 28 $^oC$ for 4-6 hr and different volumes of it were plated on antibiotic containing YEB plates. Transformed colonies were checked by colony PCR, plasmid isolation and restriction digestion.

### 3.7: Transformation of *Nicotiana tabaccum* cv Petit Havana with *A. tumifaciens*

Tobacco leaf discs (*Nicotiana tabaccum* cv Petit Havana) were transformed with either pCFNcb or pCFNcb* containing constructs independently as described before (Horsch et al., 1985).

Primary culture of *A. tumefaciens* was initiated from a single colony in 5 ml YEB containing rifampicin (25 mg $l^{-1}$), gentamycin (40 mg $l^{-1}$) and kanamycin (50 mg $l^{-1}$) at 28C with vigorous shaking. After 2days, 500 µl of the primary culture was used to initiate 5 ml of secondary culture. One ml of the secondary culture was inoculated in 30 ml YEB (without the antibiotic) and grown for 3-4 hr till an $O.D_{600}$ reached to 0.6. Cells were pelleted at 3,500 rpm for 10 min at RT. The pellet was resuspended in 5 ml of Murashige-Skoog (MS) medium. The leaf discs were co-infected with *Agrobacterium tumefaciens* at a final $O.D_{600}$ 0.3 for 20 min. The co-infected leaf discs were co-cultured in MS medium supplemented with benzyladenine purine (1.0 mg $l^{-1}$) and napthylacetic acid (0.1 mg $l^{-1}$) ($B_1N_{0.1}$) in the dark for 2 days. After 2 days, the explants were washed several times with sterile water containing augmentin (400 mg $l^{-1}$) and plated on $B_1N_{0.1}$ containing kanamycin (100 mg $l^{-1}$) and augmentin (400 mg $l^{-1}$). One shoot regenerated from each of the leaf discs were rooted on MS medium containing kanamycin (100 mg $l^{-1}$) and augmentin (400 mg $l^{-1}$).

The plants were subcultured at a regular interval of 15-20 days for 2-3 times and taken out of the medium for hardening of the plants. These plantlets were kept in sterile water in the culture tubes covered with plastic bags for 2-3 d. They were later transferred to agro-peat in plastic pots and covered with plastic bags to retain humidity. After 3-4 days, the bags were slowly opened in order to reduce humidity gradually until the plants got acclimatized. Finally, the plants were transferred to soil and grown under green house conditions for seed production.

### 3.8: Histochemical GUS assay

Gus staining was carried out using a substrate solution containing 100 mM $NaPO_4$ pH 7.0, 10 mM EDTA, 0.1% Triton X-100, 0.5 mg/ml 5-bromo-4-chloro-3-

indolyl β-D glucuronic acid (X-gluc), 2 mM each of potassium ferrocyanide and potassium ferricyanide. Whole tissue or fine sections of leaf, stem were incubated in the X-gluc containing buffer.

### 3.9: Selection of $T_1$ seeds

The $T_1$ and untransformed control seeds were surface sterilized using NaOCl for 5 mins, followed by a wash with 0.01% Triton X for 5 mins and finally with 70% Ethanol. The seeds were finally washed several times with sterile water and plated on wet Whatmann paper. Germinated seedlings were plated on MS medium containing 100 mg $l^{-1}$ kanamycin. Seedlings, which had a proper root system, were used for further studies.

### 3.10: Isolation of plant genomic DNA

Genomic DNA was isolated as described before (Murray and Thompson, 1980). Leaves were ground to a fine powder in liquid nitrogen using mortar and pestle. CTAB buffer (0.1m Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% CTAB and 0.2% *β-mercaptoethanol*) was added to the ground tissue in the ratio of 5:1. The powder was allowed to thaw at 4 $^{o}$C and transferred to Oakridge tubes. The homogenate was incubated at 65 $^{o}$C for 30-45 min. Equal volume of chloroform-isoamylalcohol (24:1) was added to the homogenate and it was mixed by gentle inversions. The whole mixture was centrifuged at 10,000 rpm for 15 min at RT and the upper aqueous phase was transferred to fresh tube. Extraction with chloroform-isoamylalcohol was repeated .To the aqueous phase 0.6 vol of isopropanol was added and incubated at room temperature for 30 min. DNA pellet was collected by centrifugation at 8,000 rpm for 20 min at 4 $^{o}$C. The pellet was air dried and dissolved in 1ml of sterile water. RNase (20 µg/ml) treatment was given at 37 $^{o}$C overnight. DNA was extracted with phenol-chloroform-isoamylalcohol (25:24:1) twice to remove proteins. DNA was re-precipitated with 2 vol ethanol and 1/10 vol of 3 M sodium acetate (pH 5.2). DNA pellet was washed with 70% ethanol, air-dried and dissolved in sterile water.

# 4: Results

## 4.1: Mapping of transcription start- and polyadenylation sites

RNA was isolated from *Petunia parodii* infected with PVCV and subjected to 5'- RACE. Adenosine at nucleotide position 6911 of the published sequence (accession number U95208) was identified as the transcription start site (Fig.1A, Ba). A consensus TATA-box is located 32 base pairs upstream of it.

Likewise, 3'-RACE revealed a polyA-site at thymidine +53, 20 nucleotides downstream of the "AATAAA" polyadenylation signal (Fig. 1A, Ba). Note, that the PVCV RNA, like that of other plant pararetroviruses and animal retroviruses, is terminally redundant and that the 5'-terminal polyadenylation site is silent (Fig. 1A).

## 4.2: Promoter Analysis

Nine different length fragments of DNA sequences upstream-, and downstream of the transcription start site (Table 1, Fig. 1A) were cloned and tested to drive transcription in a vector expressing GUS. In transfected *P. parodii* protoplasts best results were achieved with a combination of upstream sequences c and downstream sequences b (promoter FNcb; Fig. 1C). FNcb comprises 356 base pairs spanning from position -282 to +74 with a TATA box at position -32 (Fig. 1A). Inclusion of additional sequences further upstream did not increase the promoter activity and inclusion of sequences further downstream decreased the activity. Similar results were obtained with protoplasts of different origin (shown only for FNcb in Fig. 1D). Therefore, no effect of homologous chromosomal sequences as present only in the tested *P. hybrida* on the transient PVCV promoter activity has been observed.

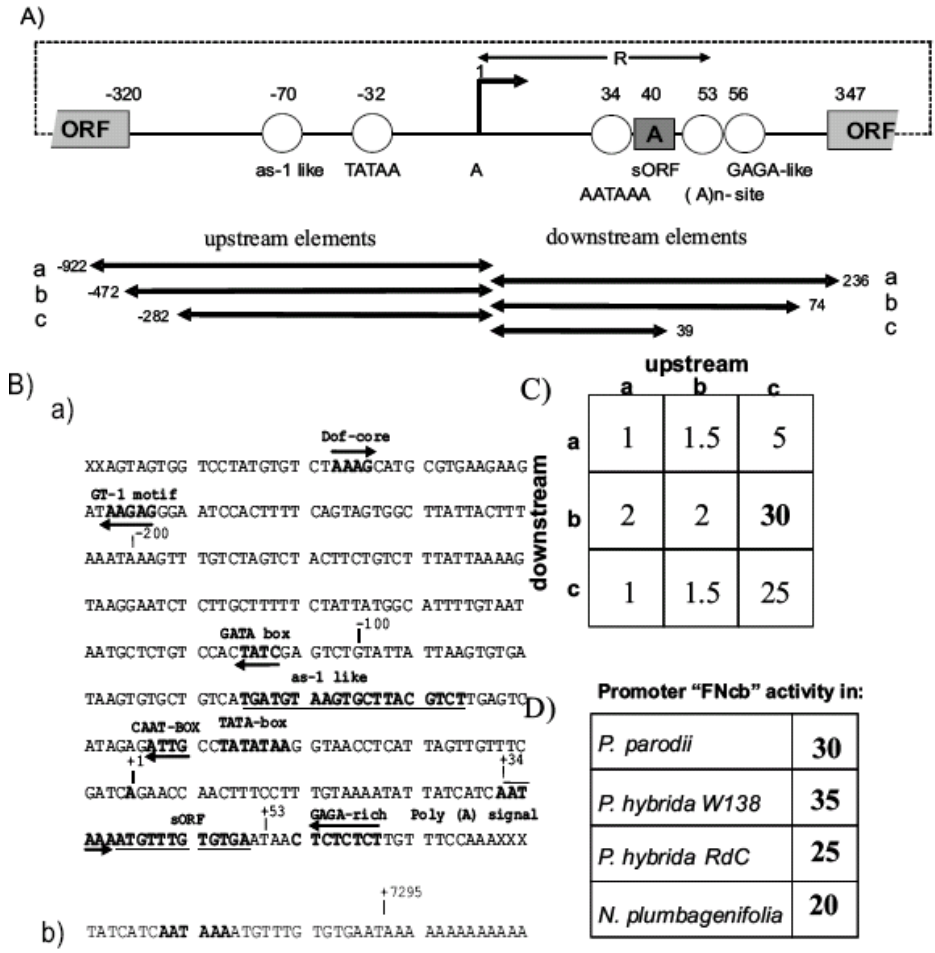


**Fig. 1**: A) Map of the promoter-region showing degenerated as-1 element, TATA-box, transcript start site, GAGA-like element, three codon sORF (A),ORF, the single large ORF of PVCV (ORF in gray), poly (A) recognition signal AATAAA polyadenylaton site (A)n-site and terminal repeats (R). a, b, c are the positions of the downstream- and upstream sequences tested in combination for promoter activity, B a) Sequence of 356bp FNcb showing regulatory elements on DNA b) poly (A) site on mRNA. The (A)n site is over-read on the circular viral DNA at position 53 and recognized at the second accounter at postion 7259, giving rise the terminal repeats (R). C) Promoter activities of combinations of upstream and downstream elements in % of CaMV 35S promoter activity. GUS was used as reporter. Errors are about +/- 10%. D) Activity of promoter "FNcb" in different plant protoplasts.

The as-1 like element with the repetitive TGACG-motifs degenerated at one base was found at positions -54 and -70 (Fig. 2Ba) upstream of the transcription start site. By reverting these putative mutations and restoring the TGACG motifs, we could in fact achieve a considerable increase in promoter activity (Fig. 2A, B). In contrast to the upstream enhancer element a putative downstream enhancer consisting of four “CT” repeats (Fig. 1Ba), suggesting a GAGA binding motif on the reverse strand, did not influence promoter activity significantly. GUS expression in presence of the CT repeat as found in construct FNcb (Fig. 1C) was only slightly increased when compared to construct FNcc (Fig.1C) in which the CT repeat was deleted.

Furthermore, in silico analysis using the PLACE database (Higo et al., 1999) showed that several putative cis-regulatory on the forward strand (Dof-core) as well as the reverse strand (GT-1, GATA-box and GAGA-rich domains) exist in the PVCV promoter construct (FNcb; Fig. 1Ba).

A)

| Promoter | as-1 like sequence | Position |
|---|---|---|
| CaMV | **TGACG**TAAGGGA**TGACG** | -79 to -63 |
| CVMV | **AGACG**TAAGCAC**TGACG** | -219 to -203 |
| CoYMV | **TGATG**ATGTCAT**TGATG** | -222 to -206 |
| FMV | **TGACG**AACGCAG**TGACG** | -73 to -57 |
| PVCV (FNcb) | **TGATG**TAAGTGC**TTACG** | -70 to -54 |
| PVCV (FNcb*) | **TGACG**TAAGTGC**TGACG** | -70 to -54 |

B)

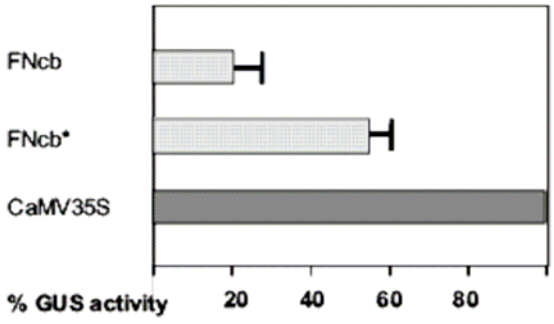


**Fig. 2:** Effect of mutation in as-1 like element of PVCV promoter . A) sequence comparison of as-1 like elements present in promoters of different plant pararetroviruses with CaMV as-1 element. B) FNcb activity can be improved considerably by restoring the degenerated repetitive sequence of the as-1 like element FNcb*. Promoter activities in % of CaMV 35S promoter activity. Error are about +/- 10%.

### 4.3: Tissue specificity of the PVCV promoter in transgenic tobacco

10-15 primary transformants (T0) were regenerated from transformed leaf discs selecting one shoot per callus produced. Histochemical analysis for GUS activity using an X-GlcA staining solution showed thorough coloration of calli and shoots derived from tissue transformed with the PVCV promoter construct (pCFNcb or pCFNcb*) as well as from tissue transformed with the 35S promoter of CaMV. Untransformed control tissue did not showed staining. The GUS staining was repeated on the material after root production. Whereas for the FNcb and 35S promoters no tissue specificity could be detected and GUS staining in leaves, stem and root parts occurred. The FNcb* promoter showed a preference for leaf and stem tissue but seemingly lacked activity in roots. Transverse sections of stem and leaf showed that all promoters were active in the epidermis and vascular bundles of stem tissue.

Furthermore, these plants tested positive in PCR using specific primers for the viral promoters or the GUS gene and were used for seed production to obtain stable transformed T1 generation. Around 10 T1 plants (3 months old) of 10 independent lines are growing under selective conditions at present and await further analysis (PCR, GUS staining and Southern hybridization).

## 5: Discussion

Studies of several plant pararetroviruses promoters provided insight into expression patterns and their potential for driving expression of transgene in plants. Some members of caulimoviridae family, *petunia vein clearing virus* being one of them, might be referred to as natural transgenes, since they can exist as constitutively transcribed componenets of the plant genome (Richert-Poggeler et al., 2003). The non-coding region of episomal and chromosomal virus forms display 95% homology (Richert-Poggeler et al., 2003). We therefore analyzed the non-coding region from an infectious clone of PVCV (Gene Bank accession No. U95208) for regulatory sequences involved in transcription and determined successfully promoter elements, transcription start site and polyadenylation signal together with polyadenylation site that were preserved in the integrated

sequences (Gene Bank accession No. AY228106) of this endogenous plant pararetrovirus.

As-1 like element are important enhancers of plant pararetroviral transcription and are generally located within the first 200 bp upstream of the promoter transcription start site (Lam and Chua, 1989a; Sanger et al., 1990; Medberry et al., 1992; Verdaguer et al., 1996; Harper and Hull, 1998) or further upstream in case of the *banana streak virus* promoter (Remans et al., 2005). However, the nucleotide composition of these elements can be variable and cause degeneration of consensus sequence consisting of two 8 bp palindromes (TGACGTCA) spaced by 4 nucleotides in case of the caulimoviruses (Krawczyk et al., 2002). Our studies revealed that the PVCV promoter carring a highly degenerated as-1 like element is still capable of driving transcription successfully and restoring the first half of the two palindromes (TGACG) to the sequence found in CaMV significantly increased promoter activity probably due to improved recognition by TGA transcription factors (Lam et al., 1989b; Krawczyk et al., 2002). We could recently show (Noreen et al., 2005, submitted), that the non-coding region of PVCV is hypomethylated. We assume that the selective advantage for the observed mutations within the PVCV as-1 like element is a reduction in methylation targets (cytosine residue) recognized by the host silencing machinery. However, for CaMV the importance of cytosine residues for the as-1 element function has been shown previously (Hohn et al., 1996) and might explain the reduce activity of the PVCV promoter.

The function of an upstream enhancer element consisting of CT repeats that in reverse orientation are known as GAGA elements and interact with transcription factors, as reported for CaMV and RTBV (He et al., 2002; Pauli et al., 2004) was not observed for the PVCV promoter. However, the motif was present only in a single version in the PVCV promoter constructs and not duplicated as in CaMV (Pauli et al., 2004) or of multiple locations as in RTBV (He et al., 2002) which may not be sufficient for the corresponding transcription factor to bind.

The identified PVCV promoter was further analyzed for the putative plant transcription factor binding sites using PLACE database of plant cis-acting regulatory DNA elements (Higo et al., 1999). This analysis revealed that like other plant pararetroviral promoters (Remans et al., 2005), the PVCV promoter has cis-acting elements proximal to the transcription start site. These putative elements are both upstream and downstream of transcription start site and are conserved in both episomal and chromosomal forms of PVCV (Gene Bank accession No. U95208 and Ay228106) indicating multiple interactions of the virus DNA with the host cell transcriptional machinery. For example, similarity to GT-1 elements that are known to be involved in activation of transcription through direct interaction with TATA binding protein TFIIA in Arabidopsis (Le Gourrierec et al., 1999) were found in the PVCV promoter. The impact of putative recognition sequences for plant transcription factors like “Dof core” and “GATA box” on PVCV transcription still has to be experimentally evaluated.

Preliminary results regarding PVCV promoter activity in primary transformants of tobacco indicate constitutive and similar tissue expression observed for the 35S promoter of CaMV. However, it will be important to follow the expression patterns through further successive generations to confirm and determine the stability of expression.

## 6: References


**Benfey, P.N., Ren, L., and Chua, N.H.** (1990). Combinatorial and synergistic properties of CaMV 35S enhancer subdomains. Embo J **9,** 1685-1696.

**Bhattacharyya-Pakrasi, M., Peng, J., Elmer, J.S., Laco, G., Shen, P., Kaniewska, M.B., Kononowicz, H., Wen, F., Hodges, T.K., and Beachy, R.N.** (1993). Specificity of a promoter from the rice tungro bacilliform virus for expression in phloem tissues. Plant J **4,** 71-79.

**Bouchez, D., Tokuhisa, J.G., Llewellyn, D.J., Dennis, E.S., and Ellis, J.G.** (1989). The ocs-element is a component of the promoters of several T-DNA and plant viral genes. Embo J **8,** 4197-4204.

**Fromm, H., Katagiri, F., and Chua, N.H.** (1989). An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts. Plant Cell **1,** 977-984.

**Glasheen, B.M., Polashock, J.J., Lawrence, D.M., Gillett, J.M., Ramsdell, D.C., Vorsa, N., and Hillman, B.I.** (2002). Cloning, sequencing, and promoter identification of Blueberry red ringspot virus, a member of the family Caulimoviridae with similarities to the "Soybean chlorotic mottle-like" genus. Arch Virol **147,** 2169-2186.

**Goodall, G.J., Wiebauer, K., and Filipowicz, W.** (1990). Analysis of pre-mRNA processing in transfected plant protoplasts. Methods Enzymol **181,** 148-161.

**Harper, G., and Hull, R.** (1998). Cloning and sequence analysis of banana streak virus DNA. Virus Genes **17,** 271-278.

**He, X., Futterer, J., and Hohn, T.** (2002). Contribution of downstream promoter elements to transcriptional regulation of the rice tungro bacilliform virus promoter. Nucleic Acids Res **30,** 497-506.

**Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T.** (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res **27,** 297-300.

**Hohn, T., and Richert-Poggeler, K.R.** (2005, submitted). Replication of plant pararetroviruses. Recent advances in DNA virus replication. Research Singpost transworld (publisher).

**Hohn, T., Corsten, S., Rieke, S., Muller, M., and Rothnie, H.** (1996). Methylation of coding region alone inhibits gene expression in plant protoplasts. Proc Natl Acad Sci U S A **93,** 8334-8339.

**Horsch, R.B., Rogers, S.G., and Fraley, R.T.** (1985). Transgenic plants. Cold Spring Harb Symp Quant Biol **50,** 433-437.

**Jefferson, R.A., Burgess, S.M., and Hirsh, D.** (1986). beta-Glucuronidase from Escherichia coli as a gene-fusion marker. Proc Natl Acad Sci U S A **83,** 8447-8451.

**Katagiri, F., Lam, E., and Chua, N.H.** (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature **340,** 727-730.

**Kloti, A., Henrich, C., Bieri, S., He, X., Chen, G., Burkhardt, P.K., Wunn, J., Lucca, P., Hohn, T., Potrykus, I., and Futterer, J.** (1999). Upstream and downstream sequence elements determine the specificity of the rice tungro bacilliform virus promoter and influence RNA production after transcription initiation. Plant Mol Biol **40,** 249-266.

**Krawczyk, S., Thurow, C., Niggeweg, R., and Gatz, C.** (2002). Analysis of the spacing between the two palindromes of activation sequence-1 with respect to binding to different TGA factors and transcriptional activation potential. Nucleic Acids Res **30,** 775-781.

**Lam, E., and Chua, N.H.** (1989a). ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in Cab promoters. Plant Cell **1,** 1147-1156.

**Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.X., and Chua, N.H.** (1989). Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. Proc Natl Acad Sci U S A **86,** 7890-7894.

**Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.X., and Chua, N.H.** (1989b). Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. Proc Natl Acad Sci U S A **86,** 7890-7894.

**Le Gourrierec, J., Li, Y.F., and Zhou, D.X.** (1999). Transcriptional activation by Arabidopsis GT-1 may be through interaction with TFIIA-TBP-TATA complex. Plant J **18,** 663-668.

**Medberry, S.L., Lockhart, B.E., and Olszewski, N.E.** (1992). The Commelina yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues. Plant Cell **4,** 185-192.

**Murray, M.G., and Thompson, W.F.** (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res **8,** 4321-4325.

**Niggeweg, R., Thurow, C., Kegler, C., and Gatz, C.** (2000). Tobacco transcription factor TGA2.2 is the main component of as-1-binding factor ASF-1 and is involved in salicylic acid- and auxin-inducible expression of as-1-containing target promoters. J Biol Chem **275,** 19897-19905.

**Noreen, F., Akbergenov, R., Hohn, T., and Richert-Poggeler, K.R.** (2005, submitted). Different silencing states of endogenous *petunia vein clearing virus* in two distinct petunia hybrida lines.

**Odell, J.T., Nagy, F., and Chua, N.H.** (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature **313,** 810-812.

**Olszewski, N., Hagen, G., and Guilfoyle, T.J.** (1982). A transcriptionally active, covalently closed minichromosome of cauliflower mosaic virus DNA isolated from infected turnip leaves. Cell **29,** 395-402.

**Pauli, S., Rothnie, H.M., Chen, G., He, X., and Hohn, T.** (2004). The cauliflower mosaic virus 35S promoter extends into the transcribed region. J Virol **78,** 12120-12128.

**Pooggin, M.M., Hohn, T., and Futterer, J.** (2000). Role of a short open reading frame in ribosome shunt on the cauliflower mosaic virus RNA leader. J Biol Chem **275,** 17288-17296.

**Remans, T., Grof, C.P., Ebert, P.R., and Schenk, P.M.** (2005). Functional promoter analysis using an approach based on an in vitro evolution strategy. Biotechniques **38,** 209-210, 212, 214-206.

**Richert-Poggeler, K.R., and Shepherd, R.J.** (1997). Petunia vein-clearing virus: a plant pararetrovirus with the core sequences for an integrase function. Virology **236,** 137-146.

**Richert-Poggeler, K.R., Noreen, F., Schwarzacher, T., Harper, G., and Hohn, T.** (2003). Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. Embo J **22,** 4836-4845.

**Rothnie, H.M., Chapdelaine, Y., and Hohn, T.** (1994). Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv Virus Res **44,** 1-67.

**Sanger, M., Daubert, S., and Goodman, R.M.** (1990). Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. Plant Mol Biol **14,** 433-443.

**Schenk, P.M., Remans, T., Sagi, L., Elliott, A.R., Dietzgen, R.G., Swennen, R., Ebert, P.R., Grof, C.P., and Manners, J.M.** (2001). Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants. Plant Mol Biol **47,** 399-412.

**Stavolone, L., Kononova, M., Pauli, S., Ragozzino, A., de Haan, P., Milligan, S., Lawton, K., and Hohn, T.** (2003). Cestrum yellow leaf curling virus (CmYLCV) promoter: a new strong constitutive promoter for heterologous gene expression in a wide variety of crops. Plant Mol Biol **53,** 663-673.

**Verdaguer, B., de Kochko, A., Beachy, R.N., and Fauquet, C.** (1996). Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Mol Biol **31,** 1129-1139.

**Yang, I.C., Iommarini, J.P., Becker, D.K., Hafner, G.J., Dale, J.L., and Harding, R.M.** (2003). A promoter derived from taro bacilliform badnavirus

drives strong expression in transgenic banana and tobacco plants. Plant Cell Rep **21,** 1199-1206.

.

# Chapter 4

## RNAi-mediated resistance targeting to pararetrovirus in banana plants

Faiza Noreen[1], Marie-Line[2], Thomas Hohn[1], Katja Richert-Pöggeler[1]
[1]University of Basel, Botanical Institute, Plant Health Unit, Schönbeinstrasse 6, 4056 Basel, Switzerland. [2]CIRAD Montpellier, France.

(Manuscript in preparation)

## 1: Introduction

Gene silencing in plants occurs at both transcriptional (TGS) and posttranscriptional (PTGS) levels (Fagard and Vaucheret, 2000a). Double stranded RNA (dsRNA), which is processed to short interfering RNAs (siRNA) of 21-24 nucleotides, plays a key role as an inducer in both silencing processes. The siRNAs corresponding to promoter sequences direct the silencing machinery to block transcription of homologous promoters (TGS), whereas those corresponding to transcribed sequences direct the silencing machinery to degrade homologous RNAs (PTGS) (Matzke et al., 2004). TGS accompanied by chromatin methylation and remodeling and PTGS associated with RNA degradation can neutralize endogenous and exogenous invaders such as transgene, viruses and transposons (Baulcombe, 2004; Matzke and Birchler, 2005).

The concept of obtaining resistance to viruses by transformation with genes derived from the pathogen's genome is a well-exploited and highly effective procedure to fight viruses as causal agents of diseases in plants (Fitchen and Beachy, 1993). The essence of RNAi technology is the delivery of dsRNA as a potent activator of RNA silencing into an organism, or cell, with the purpose of triggering sequence-specific degradation of homologous target RNAs. During recent years, dsRNAs or self-complementary hairpin RNAs that are transcribed from engineered inverted repeats were shown to be potent inducers of a gene silencing response when directed against trangenes (Waterhouse et al., 1998; Johansen and Carrington, 2001). Furthermore, plants transformed with constructs that produce RNAs capable of duplex formation induced virus immunity with almost 100% efficiency when targeted against viruses (Smith et al., 2000; Kalantidis et al., 2002). The major technical limitation for this technology is that many important plant crop species are difficult to transform. Recently it has been demonstrated that RNAi can be successfully triggered against plant viruses by transient expression from a plasmid vector containing an inverted repeat of the target promoter sequence (Pooggin et al., 2003; Tenllado et al., 2004). This produces promoter dsRNAs, to be processed by dicer into

short RNAs. These siRNA signals target methylation to the homologous promoter (endogenous or transgene) by one or more DNA methyltransferases and heterochromatization by chromatin factors (Matzke et al., 2004). As a result multiplication and spread of homologus highly replicating plant viruses are effectively blocked.

One of the major constraints to the international breeding efforts of banana has been the Obino l' Ewai (AAB group) isolate of banana streak virus named as *Banana streak Obino l'Ewai virus* (BSOEV; (Geering et al., 2005). The genome of BSOEV is typical of badnaviruses, containing three open reading frames (ORFs). BSOEV is particularly common in new *Musa acuminata* (A genome) X *Musa balbisiana* (B genome) hybrids created by banana breeding programs (Harper et al., 1999). This infection is believed to arise *de novo* from viral sequences integrated in the host genome. These integrants were shown to be present in various BB and AAB genotypes but not in the AAA genotype of Cavendish banana, the most important banana in world trade, suggesting that integrated BSOEV are associated with the B genome of *Musa* (Geering et al., 2005). Integrated sequences have been shown to be activated during *in vitro* propagation and possibly other stress factors (Harper et al., 1999). This phenomenon has prevented the deployment of tissue cultures of improved banana hybrids.

As vegetative propagation appears to play a large role in transmission of BSOEV, the most effective means to control the disease is to ensure that source plants used for propagation are virus-free. Alternatively, treatments that would inhibit the activation of integrated viral genomes could prevent the formation of episomal virus during propagation. Our effort is to develop RNA-mediated banana streak virus resistance via TGS and/or PTGS. Plants transiently expressing dsRNA homologous to the BSOEV promoter are expected to inhibit or reduce the activation of viral sequences from both endogenous and episomal sources.

## 2: Material and Methods

**2.1: Construction of a full-length BSOEV clone and proof of infectivity (to be provided by collaborators from CIRAD, Montpellier)**

**2.2: Generation of BSOEV infected banana plants**

For the studies *Musa acuminata* (A genome) plants from the triploid cultivar "Cavendish" were used. Young Cavendish plants were infected with the infectious full-length clone of BSOEV using agroinoculation through infection of leaves or tissue surrounding meristematic cells. Successful BSOEV infection was confirmed using IC-PCR. Infected banana plantlets were multiplied via *in vitro* culture by Vitropic (Montpellier, France). After propagation under sterile conditions 80 plants were transferred to soil for the RNAi studies.

**2.3: Cloning of inverted repeats**

We developed constructs containing the BSOEV promoter of different length in sense and antisense orientation seperated by a synthetic plant intron (Goodall and Filipowicz, 1989) according to the strategy described by Pooggin et al., (2003). Constructs were designed to express, under the control of the cauliflower mosaic virus 35S promoter and poly (A) signal, dsRNA in a hairpin formation cognate to the bi-directional promoter of BSOEV. The following recombinant plasmids were tested: BSOEV-short hairpin (Fig. 2A: hp-s, 323-bp long from nucleotide position 6956-7279 within the viral genome) and BSOEV-long hairpin (hp-l, 474-bp long from nucleotide position 6956 to 64 on the circular DNA genome, Accession No.NC_003381). For cloning of the selected BSOEV promoter regions in sense orientation the restriction sites of *Kpn*I and *Cla*I and for cloning in antisense orientation the restriction sites of *Bam*HI and *Xho*I were used.

**2.4: DNA extraction and sample preparation**

Crude extracts were prepared from BSOEV infected Cavendish leaf tissues for IC-PCR. 0.5 g of leaves per plant were grinded in 5 ml PBS-buffer containing 2% polyvinylpyrrolidine, 0.2% sodium sulphite and 0.2% BSA, followed by centrifugation for 5 min at 4°C.

**2.5: Immunocapture PCR (IC-PCR)**

Sterile 0.5 ml polypropylene microcentrifuge tubes were coated with 25 µl of IgG (2ng/ml) purified from polyclonal rabbit antiserum (provided by B. E. L. Lockhart), in carbonate buffer, and incubated at 37°C for 2h according to the protocol of (Harper et al., 1999). Unbound antibodies were removed in three washes with PBS-Tween (PBS-buffer with 0.05% Tween-20) and 25 µl of prepared plant extract (see above) were loaded and incubated overnight at 4°C. After incubation the tubes were washed with PBS-Tween three times followed by two times washing with sterile water and dried briefly. PCR was carried out in the same tubes using BSOEV-specific primers:RD-Forward:ATCTGAAGGTGTGTT GATCAATGC (nt position in BSOEV 6499-6522), RD-Reverse:GCTCACTCCGC ATCTTATC (nt position in BSOEV 7020-6998). PCR conditions included initial denaturation at 94°C for 5 min followed by 35 cycles consisting of the following steps: denaturation at 94°C for 30 sec, annealing at 64°C for 30 sec and extension at 72°C for 1 min.

**2.6: Particle bombardment**

Biolistic inoculation of the hairpin constructs was performed using a homemade gun device. For each series of five shots, 5 µg of recombinant plasmid DNA was precipitated on tungsten particles. Macrocarrier, rupture discs and stopping screens were sterilized with 70% ethanol (EtOH). Tungsten particles (60 mg) were washed with EtOH and sterile water and resuspended in 1ml of 50% glycerol (sterile). For macro-carrier preparation the protocol as described earlier (Richert-Poggeler et al., 2003) was used. 54 two weeks old BSOEV infected plants harboring three true leaves were used for particle bombardment. As a control 5-10 plants were bombarded either with DNA-free macrocarriers or with a plasmid vector bearing BSOEV non homologous sequences of the bidirectional promoter of *Vigna mungo yellow mosaic virus* (VMYMV, Pooggin et al., 2003). Leaf samples were collected from leaves directly before and in weekly intervals after bombardment and were analyzed by IC-PCR.

## 3: Results

The potential of RNA-mediated banana streak virus resistance to prevent the outbreak of virus infection upon rare spontaneous induction of endogenous BSOEV in tissue culture was tested in a model system. The generated BSOEV full-length clone proved to be infectious after agroinoculation of Cavendish (AAA genome) banana plants that do not contain endogenous BSOEV sequences (data not shown).

BSOEV infection mostly persisted during micropropagation of the Cavendish banana plants using tissue culture. From a typical experiment, 80 plants could be obtained, 75% of which were infected, as shown by IC-PCR (Fig.1).

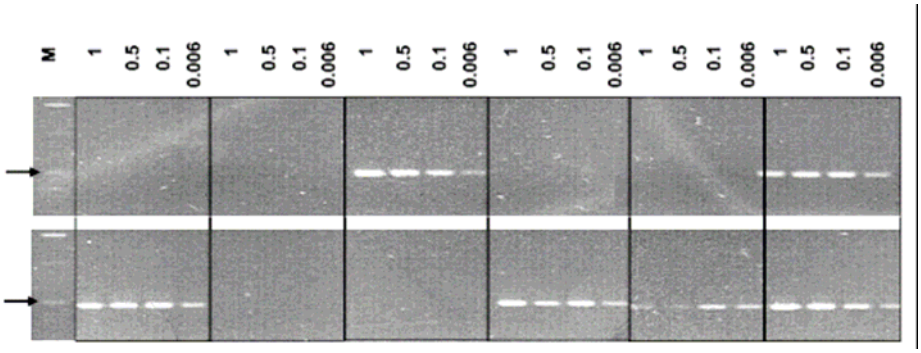


**Fig 1:** Specific identification of episomal BSOEV using IC-PCR. A screen of 12 plants is shown with 4 different dilutions of each sample. Gels are ethidium bromide stained of amplified fragment obtained using specific BSOEV primers. Arrows indicate specific amplification of a 520-bp fragment of BSOEV.

These plants were bombarded with either the short (hp-s) or long (hp-l) version of the BSOEV hairpin construct (Fig. 2) to transiently produce dsRNA molecules homologous to BSOEV promoter sequences.

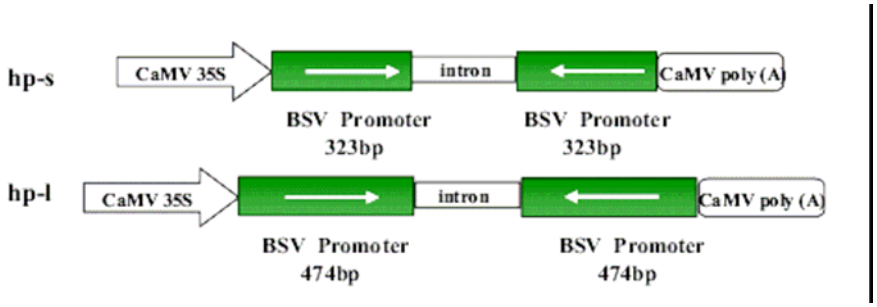


**Fig 2:** Interfering DNA construct of two different lengths hairpin short (hp-s) 323bp fragment, hairpin long (hp-l) 474pb.

As parameter for effects on infection by the dsDNA virus, relative DNA levels were measured before and after bombardment with the constructs using IC-PCR, and only those plantlets that originally contained measurable virus DNA quantities were considered in Fig. 3A. Plants could be grouped according to their response to bombardment. About one third of the plants accumulated considerable amounts of viral DNA, much like in the untreated control (Fig. 3B). We assume that these plants escaped delivery of hairpin RNA or, alternatively in these plants silencing was suppressed. A second group showed recovery from virus infection still lasting to 4 weeks after bombardment. We assume that in this case of “full recovery”, virus silencing was effective. In a third group, observed after use of the shorter hairpin construct, plants recovered at three weeks but the virus titer raised again after 4 weeks. We assume that in this case of “transient recovery”, a low level of viruses could escape silencing and managed to produce enough silencing suppressor to resume virus replication.

A)

| | 1week before | 1 week after | 2 weeks after | 3 weeks after | 4 weeks after | numbers | | Result |
|---|---|---|---|---|---|---|---|---|
| hp-I | ++++ | +++ | +++ | - | - | 1 | 4/12 | Full Recovery |
| | ++++ | ++++ | ++ | ++ | - | 2 | | |
| | ++++ | ++ | + | + | - | 1 | | |
| | ++++ | +++ | ++ | ++ | ++ | 2 | 8/12 | No Recovery |
| | ++++ | + | + | ++ | ++ | 2 | | |
| | ++++ | ++++ | ++++ | ++++ | ++++ | 4 | | |
| Control | ++++ | ++++ | ++++ | ++++ | ++++ | 3 | 3/3 | |
| hp-s | ++++ | +++ | +++ | - | - | 1 | 4/15 | Full Recovery |
| | ++++ | ++ | +++ | +++ | - | 2 | | |
| | ++++ | + | + | - | - | 1 | | |
| | ++++ | + | + | + | + | 4 | 9/15 | No Recovery |
| | ++++ | ++++ | ++++ | ++++ | ++++ | 5 | | |
| | ++++ | ++ | + | | ++++ | 2 | 2/15 | Transient Recovery |

B)

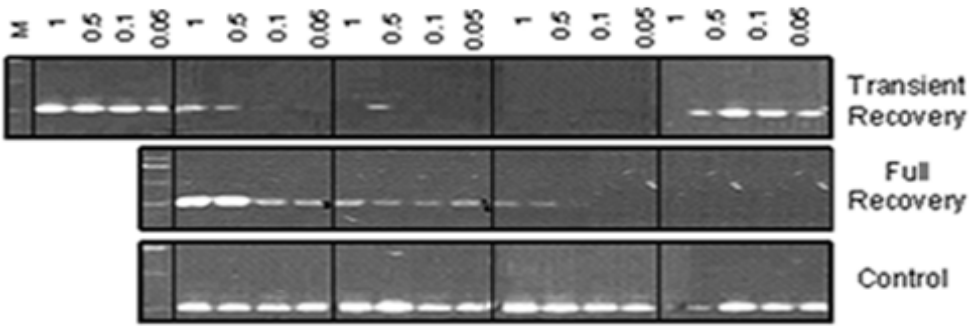


**Fig 3.** Effect of bombardment of BSV-OI infected 2 week-old seedlings with hairpin constructs hp-l and hp-s. A) Results obtained as described in M&M from semiquantitative IC-PCR from different dilutions of RNA extracts before and after bombardment with primers to cover the promoter region. The number of "+" reflects relative concentrations of viral DNA. B) Examples of IC-PCR gels for "transient recovery", "full recovery" and "control". No recovery appeared similar than the control.

## 3: Discussion and outlook

The applied model system could demonstrate that transient application of RNAi has the potential to install long term protection against BSOEV infection. Furthermore preliminary evidence could be provided that the success rate of inhibition depends of the existing virus titer in treated plants. So far whole leaves were bombarded and the possibility exists that all infected cells received the silencing construct and therefore viral infection persisted as shown in cases of no recovery and transient recovery (Fig. 3B). Therefore, it may be advisable to repeat the biolistic inoculation with the silencing construct after two weeks to inhibit the observed transient recovery.

Caulimoviridae life cycle happens in both nuclear and cytoplasmic compartments of the plant cell offering two locations of attack for an RNAi based defense system.

The designed inverted repeat of the delivered plasmid construct recognizes specifically BSOEV promoter sequences that are present on viral DNA and the terminally redundant RNA molecules. The reduction or absence of viral DNA in the successfully treated banana indicates that DNA replication is impaired. Since caulimoviridae are retroelements they use reverse transcritption of mRNA for DNA synthesis. Therefore we assume viral mRNA production is targeted in the following ways: the produced dsRNA of the silencing construct

gets cleaved into siRNAs from nuclear dicers and interacts with the viral minichromosomes present in the nucleus inhibiting transcription by DNA methylation of the promoter sequences. Secondly, the generated siRNAs could be transported to the cytoplasm and used for priming dsRNA synthesis on viral mRNA molecules resulting in mRNA degradation and the production of new siRNAs.

The loss or reduction of viral DNA in non-bombarded, upper leaves provides indirect evidence for the presence of a mobile silencing signal typically associated with PTGS (Fagard and Vaucheret, 2000b).

The promising results obtained so far prompt us to employ this strategy of transient RNAi further to “immunize” calli derived from banana hybrids that carry potentially pathogenic BSOEV sequences in their genome. The following advantages are inherent in such an application: i) inhibition of viral escape at an earliest time point possible during micropropagation, ii) less cells compared to leaf tissue have to be targeted and therefore the efficiency of RNAi construct delivery to each cell increases and iii) long term protection.

## 4: References


**Baulcombe, D.** (2004). RNA silencing in plants. Nature **431,** 356-363.

**Fagard, M., and Vaucheret, H.** (2000a). (TRANS)GENE SILENCING IN PLANTS: How Many Mechanisms? Annu Rev Plant Physiol Plant Mol Biol **51,** 167-194.

**Fagard, M., and Vaucheret, H.** (2000b). Systemic silencing signal(s). Plant Mol Biol **43,** 285-293.

**Fitchen, J.H., and Beachy, R.N.** (1993). Genetically engineered protection against viruses in transgenic plants. Annu Rev Microbiol **47,** 739-763.

**Geering, A.D., Olszewski, N.E., Harper, G., Lockhart, B.E., Hull, R., and Thomas, J.E.** (2005). Banana contains a diverse array of endogenous badnaviruses. J Gen Virol **86,** 511-520.

**Goodall, G.J., and Filipowicz, W.** (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58,** 473-483.

**Harper, G., Dahal, G., Thottappilly, G., and Hull, R.** (1999). Detection of episomal banana streak badnavirus by IC-PCR. J Virol Methods **79,** 1-8.

**Johansen, L.K., and Carrington, J.C.** (2001). Silencing on the spot. Induction and suppression of RNA silencing in the Agrobacterium-mediated transient expression system. Plant Physiol **126,** 930-938.

**Kalantidis, K., Psaradakis, S., Tabler, M., and Tsagris, M.** (2002). The occurrence of CMV-specific short Rnas in transgenic tobacco expressing virus-derived double-stranded RNA is indicative of resistance to the virus. Mol Plant Microbe Interact **15,** 826-833.

**Matzke, M., Aufsatz, W., Kanno, T., Daxinger, L., Papp, I., Mette, M.F., and Matzke, A.J.** (2004). Genetic analysis of RNA-mediated transcriptional gene silencing. Biochim Biophys Acta **1677,** 129-141.

**Matzke, M.A., and Birchler, J.A.** (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet **6,** 24-35.

**Pooggin, M., Shivaprasad, P.V., Veluthambi, K., and Hohn, T.** (2003). RNAi targeting of DNA virus in plants. Nat Biotechnol **21,** 131-132.

**Richert-Poggeler, K.R., Noreen, F., Schwarzacher, T., Harper, G., and Hohn, T.** (2003). Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. Embo J **22,** 4836-4845.

**Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M.** (2000). Total silencing by intron-spliced hairpin RNAs. Nature **407,** 319-320.

**Tenllado, F., Llave, C., and Diaz-Ruiz, J.R.** (2004). RNA interference as a new biotechnological tool for the control of virus diseases in plants. Virus Res **102,** 85-96.

**Waterhouse, P.M., Graham, M.W., and Wang, M.B.** (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc Natl Acad Sci U S A **95,** 13959-13964.

# Chapter 5

## Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia

Katja Richert-Pöggeler[1*],Faiza Noreen[1], Trude Schwarzacher[2], Glyn Harper[3] and Thomas Hohn[4],
[1]Friedrich Miescher Institute, PO Box 2543, CH-4002 Basel, Switzerland, [2] University of Leicester, Leicester LE1 7RH and [3] John Innes Centre, Colney Lane, Norwich NR4 7UH, UK [4] Present address: University of Basel, Botanical Institute, Plant Health Unit, Schönbeinstrasse 6, CH-4056, Basel, Switzerland [*] Corresponding author

(The EMBO Journal 2003, 22: 4836-45)

## 1: Abstract

Infection by an endogenous pararetrovirus using forms of both episomal and chromosomal origin has been demonstrated and characterized, together with evidence that petunia vein clearing virus (PVCV) is a constituent of the *Petunia hybrida* genome. Our findings allow comparative and direct analysis of horizontally and vertically transmitted virus forms and demonstrate their infectivity using biolistic transformation of a provirus-free petunia species. Some integrants within the genome of *P.hybrida* are arranged in tandem, allowing direct release of virus by transcription. In addition to known inducers of endogenous pararetroviruses, such as genome hybridization, tissue culture and abiotic stresses, we observed activation of PVCV after wounding. Our data also support the hypothesis that the host plant uses DNA methylation to control the endogenous pararetrovirus.

## 2: Results and discussion of my contribution

Cytosine methylation of endogenous PVCV sequences in uninduced and induced *P.hybrida* was examined using the isoschizomer pair *Scr*FI (does not cleave if internal C of CCWGG is methylated) and *Bst*NI (cleaves CCWGG regardless of methylation status) and a probe (PVCV-R) covering viral positions 5444−5671 (Fig 1), thus allowing detection of a variety of virus-specific fragments, including full-length copies, restriction fragments from both chromosomal and episomal forms, and integrated PVCV sequences flanked by plant DNA. Interpretation of the hybridization pattern and identification of some fragments is complicated by the high numbers of integration events (estimated at ~100−200), as well as polymorphism surrounding integration sites between PVCV copies. Nevertheless, a distinct methylation pattern was observed in plants with vein clearing symptoms compared with that in symptomless plants. In uninduced plants, undigested samples show a smear of high-molecular-weight DNA. *Spe*I digestion resolves this smear into a band of 4.8 kb, corresponding to the larger *Spe*I fragment of PVCV DNA. This band is not further digested by *Scr*FI; thus, this fragment contains no unmethylated C$^{*}$CWGG sequences.

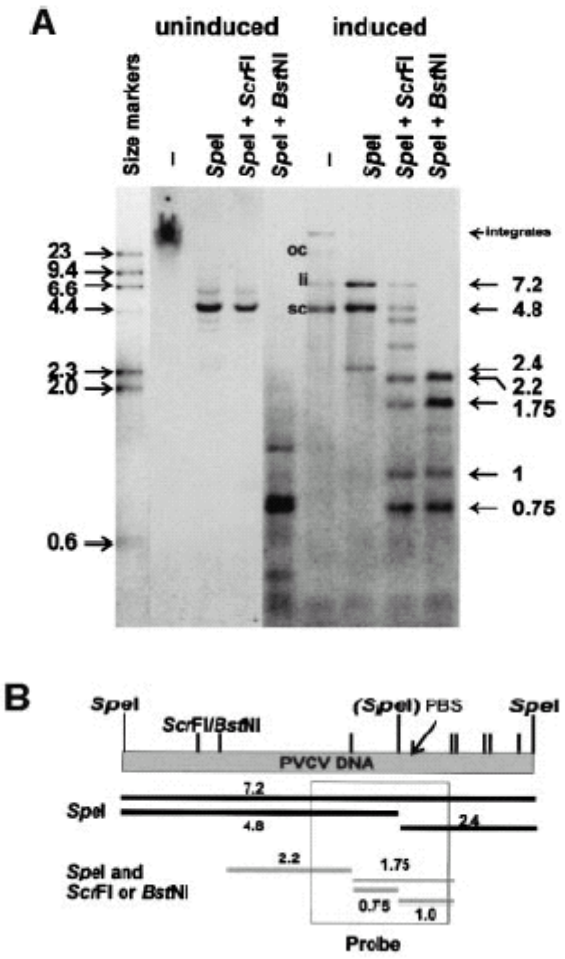


**Fig 1:** Analysis of cytosine methylation. (**A**) Genomic DNA extracted from *P.hybrida* cv. Himmelsröschen was digested with *Spe*I alone or in combination with *Scr*FI (methylation sensitive) or *Bst*NI (methylation insensitive). DNA from healthy plants (uninduced) was compared with that from symptom-expressing plants (induced). In undigested DNA from the induced sample, the positions of open circular (oc), linear (li) and supercoiled (sc) viral forms are indicated. Marker DNA consisted of a *Hin*dIII digest of digoxigenin-labelled λDNA (Roche). The sizes of major *Spe*I/*Scr*FI(*Bst*NI) fragments are shown on the right. (**B**) Linear representation of the PVCV genome, showing the positions of *Spe*I sites (in brackets, polymorphic site) and *Scr*FI/*Bst*NI sites (smaller tags). PBS, primer binding site. The region of the genome covered by probe PVCV-R is indicated, along with the sizes (kb) of the major restriction fragments expected.

In contrast, upon digestion with *Bst*NI, the 4.8 kb band disappears in favour of a collection of smaller bands. This experiment clearly shows that the bulk of the integrated PVCV DNA is methylated. Interestingly, the second (2.4 kb) *Spe*I fragment was not detected, indicating that integration and rearrangement events preferentially occur within this fragment. Upon induction, the appearance of episomal virus is clearly evident. In the undigested samples, supercoiled, linear

and (small amounts) of open circular PVCV DNA, as well as a weak band of PVCV integrates, can be discerned. Digestion with *Spe*I results in fragments of 7.2 kb, representing full-length linearized viral DNA that originated from circular forms, and 4.8/2.4 kb fragments, indicative of the polymorphism of the *Spe*I site at position 6984.

The restriction pattern after SpeI−BstNI double digestion mainly follows the cloned sequence and is different from that of the uninduced sample, indicating that only a subpopulation of integrated sequences (corresponding to the cloned sequence) was induced and/or selected. A similar pattern is obtained using the methylation-sensitive ScrFI isoschizomer. However, this digestion was incomplete, indicating some methylation of the episomal (or packaged) viral DNA. A restriction enzyme analysis using the HpaII−MspI isoschizomer pair provided similar results (data not shown).

From the frequency of occurrence (72 positive clones) in the genomic library, we estimate that ~50−100 copies of integrated viral sequences exist in the haploid petunia genome. Despite this high frequency, spontaneous release of viruses is rare. However, this release can be induced by wound stress. For retrotransposons, it has been shown that stress conditions induce transcriptional activation (Grandbastien, 1998). Wound stress would cause callus formation accompanied by chromatin replication and transient hypomethylation. PVCV sequences appear to be clustered with metaviridae-like sequence in the petunia genome, and most likely the same mechanism applys for the control of these retroelements by the host. Indeed, integrated PVCV sequences of healthy plants appear to be methylated (Fig 1). Upon induction, where only a short period of demethylation in a single plant cell at one of the ~100−200 integration loci would be enough to release the virus, a distinct pattern appears dominated by episomal forms of PVCV. Most of those are unmethylated, with the exception of a minor fraction that is methylated. Methylation of episomal virus has also been shown for CaMV (Tang and Leisner, 1998).

## Concluding remarks

This work aimed at understanding the epigenetic regulation of endogenous plant pararetroviruses. These epigenetic studies revealed mechanism used by the host to control these mobile and potentially infectious retroelements. It appears that a certain balance between silencing by the host and silencing suppression by the virus exists. The identification of such viral suppressor to counteract host control would be very helpful to understand the host-virus interaction leading to infection originated from endogenous copies.

This work showed that the silencing of endogenous petunia vein clearing virus (ePVCV-1) is impaired in the *P. hybrida* line W138 leading to distinct methylation of chromatin and suggests a partial deficiency of transcriptional gene silencing (TGS) in W138. It remains an open question to which stage of the regulatory cycle this deficiency corresponds and to answer this question further can be very useful in epigenetic field. It will be interesting to see if similar changes also occurr under inducible conditions that would lead to the release of infectious copies from the cell visible in systemic infection.

Additional to the performed studies on histone methylation, analysis of histone acetylation would provide complete picture of ePVCV-1 at histone level.

The identified PVCV promoter may be highly valuable for plant transformation system requiring heterologous promoters for gene expression to avoid TGS based on homologous sequence recognition.

Along understanding of host–virus interaction, one of the important aspects of this study was to develop strategy to fight these viruses. One of the major constraints to the international breeding efforts of banana has been the *Banana streak Obino l'Ewai virus* (BSOEV). Integrated sequences have been shown to be activated during in vitro propagation and this phenomenon has prevented the deployment of tissue culture of improved banana hybrida. Our effort was to develop RNA mediated BSOEV virus resistance via TGS and/or PTGS. Here we show for the first time in plant pararetroviruses that the application of transient RNAi has the potential to install long term protection against BSOEV infection. Preliminary results show promising effect of this

strategy which prompts us to employ this strategy of transient RNAi further to immunize calli derived from banana hybrids that carry potentially pathogenic BSOEV sequences in their genome. This will help to inhibit the viral escape at an earliest time point possible.

# Curriculum Vita

**Faiza Noreen**

E-mail: faiza.Noreen@unibas.ch

**Objectives:**

Looking forward to a bright career requiring innovative work and maximizing my capabilities with future prospects and growth, especially in Epigenetic, Biotechnology and Molecular virology. Ready to learn more.

**Education:**

- **Doctorate in Philosophy:** Friedrich Miescher Institute and botanical Institute University of Basel, Basel, Switzerland
- **Masters of philosophy (M.Phil)** in Biotechnology- 2002 from National Institute of Biotechnology and Genetic engineering (NIBGE), Faisalabad affiliated with Quaid-i -Azam University, Islamabad. With first division by securing 88% marks.
- **Master of Science (M.Sc)** In Biochemistry- 2000 from University of the Agriculture Faisalabad. With first division by securing 87% marks
- **Bachelor of Science (B.Sc**.) -1997-University of the Punjab with first division taking Chemistry, Biology and Zoology.
- **F.Sc**-1995 Board of Intermediate and Secondary Education Faisalabad taking Chemistry, Physics and Mathematics.

**Achievement/Awards:**

- Stood first in M.Phil in department of biotechnology, Quaid-i-Azam University Islamabad and awarded distinction certificate.
- Stood first in M.Sc in Faculty of Sciences, University of Agriculture Islamabad and awarded medal.
- Awarded a scholarship from European grant commission, Friedrich Miescher Institute, Switzerland.

**Research Experience:**

- PhD Project: (2002-2005) Implication of epigenetic for the control of endogenous plant *pararetroviruses*.
- M.Phil Project: (2000-2002) Characterization of pathogenecity determinants and suppressor encoded by DNA A of cotton leaf curl virus through expression from PVX vector. (Thesis)
- M.Sc Project: (1999) Infectivity analysis of geminiviruses DNA extracts in cotton and tobacco. (Thesis)

**Research Publications:**

- Noreen, F., Akbergenov, R., Hohn, T. and Richert-Pöggeler, K., R. (2005). Different heterochromatic traits of endogenous *Petunia vein clearing virus* in two distinct *Petunia hybrida* liness. The Plant Cell, submitted.

- Richert-poggler, K., Noreen, F., Schwarzacher, T. and Hohn, T. (2003) Induction of infectious petunia vein clearing (pararetro) virus from endogenous provirus in petunia. *EMBO*, 22, 4836-4845.

- Noreen, F., Deb Roy,S., Müller, M., Sarin, N, B., Hohn, T. and Richert-Pöggeler, K., R. Isolation and chrarecterization of promoter from an inducible endogenous plant pararetrovirus in petunia, Petunia vein clearing virus (PVCV) (in preparation).

- Noreen, F, Marie-Line, Hohn, T.and Richert-Pöggeler, K., R. RNAi-mediated resistance targeting to pararetrovirus in banana plants. (in preparation)

**Personal Particulars:**


| | |
|---|---|
| Name: | Faiza Noreen |
| Father name: | Muhammad yousaf |
| Sex: | Female |
| Marital Status: | Single |
| Date of Birth: | 27 November 1977 |
| Permanent Address: | 282-H-Gulistan Colony No.1 Faisalabad, Pakistan. |
| Current address | Oetlingerstrasse 10, 4057, Basel, Switzerland |
| Nationality: | Pakistani |
| Phone: | 0041-61-2673502 |
| E-mail: | faiza.noreen@unibas.ch |


**Referees:**

- Prof. Dr Thomas Hohn, University of Basel, Botanical Institute. Switzerland.
  - Thomas.hohn@fmi.ch
- Dr. Katja Richert-Poggeler, University of Basel, Botanical Institute. Switzerland.
  - K.Richert-Poeggeler@unibas.ch

- Mr. Zahid Mukhtar, Senior Scientific Officer, Plant Biotechnology Division NIBGE, Faisalabad, Pakistan.
  - zahidmukhtar@yahoo.com