



INSTITUT DE FRANCE  
Académie des sciences

# *Comptes Rendus*

---

## *Biologies*

Hervé Vaucheret

**Epigenetic management of self and non-self: lessons from 40 years of transgenic plants**

Volume 345, issue 4 (2022), p. 149-174


Published online: 1 December 2022

Issue date: 16 February 2023

<https://doi.org/10.5802/crbio.96>

**Part of Special Issue:** Insights onto Plant Biology

**Guest editors:** Michel Delseny (Laboratoire Génome et Développement des plantes, UMR5096 CNRS-Université de Perpignan Via Domitia, Université de Perpignan 66860 Perpignan, France) and Georges Pelletier (Directeur de recherche émérite à l'Institut national de la recherche agronomique, Versailles, France)

 This article is licensed under the  
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.  
<http://creativecommons.org/licenses/by/4.0/>



*Les Comptes Rendus. Biologies* sont membres du  
Centre Mersenne pour l'édition scientifique ouverte  
[www.centre-mersenne.org](http://www.centre-mersenne.org)  
e-ISSN : 1768-3238



**Keywords.** Epigenetics, Gene silencing, RNA quality control, Transgenes, Transposons, Viruses.

**Mots-clés.** Épigénétique, Extinction de gènes, Contrôle qualité des ARN, Transgènes, Transposons, Virus.

Published online: 1 December 2022, Issue date: 16 February 2023

## La version française de l'article est disponible à la suite de la version anglaise

### 1. The pre-history of gene silencing

The term gene silencing was invented when acknowledging the fact that transgenic plants do not always express the introduced transgenes. As a matter of fact, gene silencing phenomena did not appear as a consequence of transgenesis. They already existed naturally, but were not known as such. Because it would too long to list all the unexplained natural phenomena identified before the transgenic era, this introductory section will concentrate on three cases that long remained unexplained, which reflect how endogenous genes can be silenced at the transcriptional or post-transcriptional level, and how plants eliminate exogenous elements that introduce their genetic material into plant cells (Figure 1).

The most ancient description of an unstable phenotypic variability is due to Linée who described in 1749 a variant of *Linaria vulgaris* called *peloria* in which the floral symmetry is changed from bilateral to radial (Figure 1A). Because, occasionally, the variant reverts phenotypically during somatic development, it is now considered the first epigenetic variant ever described. Indeed, the floral modification is not due to a mutation but to an epigenetic modification (epimutation), in this case DNA methylation, which causes transcriptional silencing (TGS) of the *Lcyc* endogenous gene [1]. Like every epimutation, the *peloria* epimutation does not exist in a single state. Instead, a range of intermediate phenotypes exist, which intensity correlates with the degree of DNA methylation and transcriptional silencing (TGS) of the *Lcyc* endogenous gene (Figure 1A).

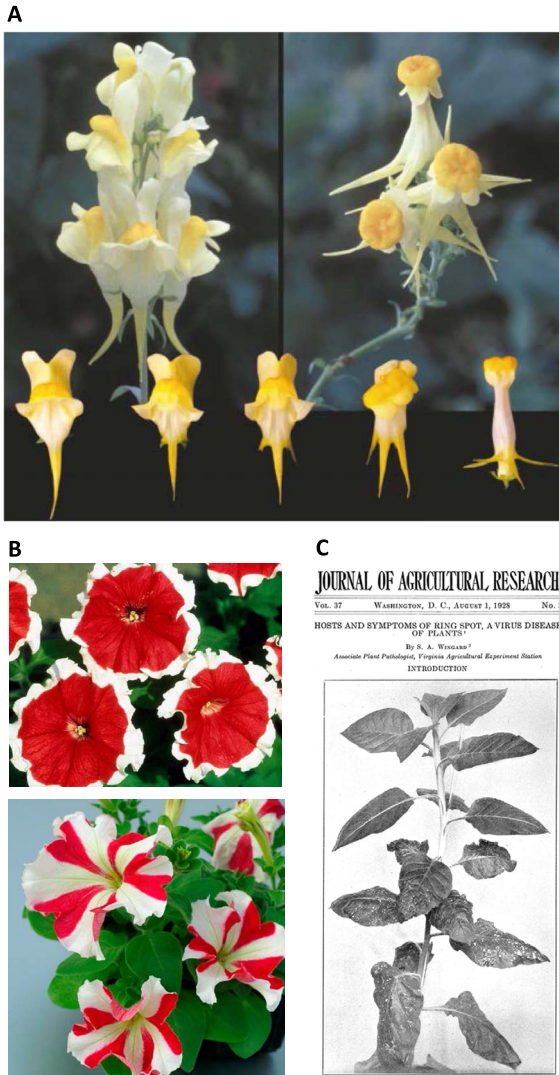
The most ancient description of a variegated phenotype dates from 1896 and relates to horticultural *Petunia hybrida* cultivars exhibiting bicolor flowers. *Petunia hybrida* is an interspecific hybrid between *Petunia axillaris*, which exhibits white flowers, and *Petunia integrifolia*, which exhibits red/purple flowers [2, 3]. Among the resulting hybrids, the Picotee cultivar displayed flowers with red centers and

white margins while the Star cultivar displayed red and white sectors alternating in a radial pattern [4] (Figure 1B). These patterns were reproduced by transgenesis when introducing extra-copies of the *CHS-A* gene, leading to the silencing of endogenous and transgenic *CHS-A* copies by PTGS [5, 6]. Eventually, the phenotype of the Picotee and Star cultivars was shown to result from endogenous duplications of the *CHS-A* gene whereas other (uniformly colored) *Petunia* varieties carry a single copy of the *CHS-A* gene [7].

The most ancient description of a gene silencing phenomenon affecting an exogenous element is due to Wingard who described in 1928 how tobacco plants can recover from virus infection and become immune to over-infection by the same virus [8] (Figure 1C). Decades were necessary to fully understand that the PTGS mechanism at work against ectopic transgene copies was also used to produce small RNAs that target viral long RNAs for destruction [9].

### 2. The discovery of transgene TGS

In the early 80's, plant biologists took advantage of the capacity of the soil bacteria *Agrobacterium tumefaciens* to transfer part of its DNA (referred to as T-DNA) to develop a universal method for plant transformation. The first genetically engineered plant was a tobacco plant carrying a modified *Agrobacterium* T-DNA in which a bacterial neomycine phosphotransferase gene and a yeast alcohol deshydrogenase gene were inserted in the T-DNA. One transgenic plant was regenerated, which transmitted the T-DNA to its self-progeny [10]. However, the *Agrobacterium nopaline synthase (NOS)* gene carried on the T-DNA was only expressed in 24% of the self-progeny (whereas it was expected to be expressed in at least 75% of the self-progeny), indicating that part of the progeny carried the T-DNA but did not express at least one of its genes. The authors found this result puzzling and unexpected but did not follow it up. In retrospect, this



**Figure 1.** Representative silencing phenomena observed in nature. (A) Flowers of wild-type *Linaria vulgaris* (top left) and of the epigenetic variant *peloria* (top right). Below are shown intermediate phenotypes observed during somatic development, ranging from wild-type (left) to full *peloria* phenotype (right). The intensity of the *peloria* phenotype correlates with the degree of methylation and transcriptional silencing (TGS) of the *Lcyc* endogenous gene. (B) Horticultural *Petunia hybrida* cultivars exhibiting PTGS of duplicated endogenous *CHS-A* genes, resulting in bicolor flowers. Top: Picotee cultivar displaying flowers with red centers and white margins.

**Figure 1. (cont.)** Bottom: Star cultivar displaying red and white sectors alternating in a radial pattern. (C) Tobacco plants recovering from virus infection through PTGS. At the bottom of the plant, the oldest leaves show severe symptoms of infection by the tobacco ringspot virus. The symptoms progressively disappear as the plant grows until it produces upper leaves that are virus-free and immune to over-infection by the same virus.

seems to be the first example of gene silencing affecting a transgenic plant reported in the literature (Antonius and Marjori Matzke, personal communication), and it is remarkable that it happened in the first transgenic plant ever obtained.

In the following years, an increasing number of reports described the obtention of transgenic plants that expressed the introduced transgenes. But still, no report of expression issues, either because no attention was paid to progeny plants that no longer expressed the transgenes or because the authors censored what could be considered negatively when trying to publish. An important step was passed when observing that sequential transformation of tobacco by a second T-DNA could lead to the reversible silencing of transgenes carried by the first T-DNA and sharing homology in their promoter sequences. Indeed, this silencing phenomenon could be reverted when segregating the second T-DNA [11]. Importantly, this first study established a correlation between DNA methylation in promoter sequences and a form of silencing that was later shown to be TGS [12]. Because silencing of the first T-DNA occurred only in the presence of the second T-DNA, this phenomenon was supposed to directly depend on copy number as was the MIP phenomenon described in the fungus *Ascobolus immersus* [13]. The study of transgenic petunia plants carrying the maize dihydrofolate reductase A1 gene first confirmed the positive correlation between copy number and silencing and methylation [14]. Further analysis revealed that an A1 transgene locus that spontaneously undergoes TGS *in cis* can trigger *in trans* the TGS of non-silenced A1 transgene locus [14], thus explaining how TGS was established during sequential transformation [11]. The study of a tobacco plant carrying multiple methylated copies of a plasmid carrying tran-

scriptionally silencing transgenes driven by the viral 19S and 35S promoters revealed that TGS triggered *in trans* depends on the homology in the promoter but not the coding sequence and is independent of the position of the target transgenes in the genome [15]. Eventually, transgene loci capable of undergoing TGS spontaneously *in cis* and subsequently *in trans* were shown to result from insertion events generating an inverted repeat structure allowing the production of double-stranded (ds)RNA from the transgene promoter sequences [16, 17]. These dsRNAs are processed into a class of short interfering (si)RNAs that are 24-nt long, which guide *de novo* DNA methylation and chromatin remodeling at homologous loci to silence them *in trans*, a process also referred to as RNA-directed DNA methylation or RdDM [18]. Various genetic screens were set up using transgenic two-component systems, allowing dissecting the mechanism underlying RdDM [18–20]. A scheme of this process is depicted in Figure 2.

### 3. Natural TGS phenomena

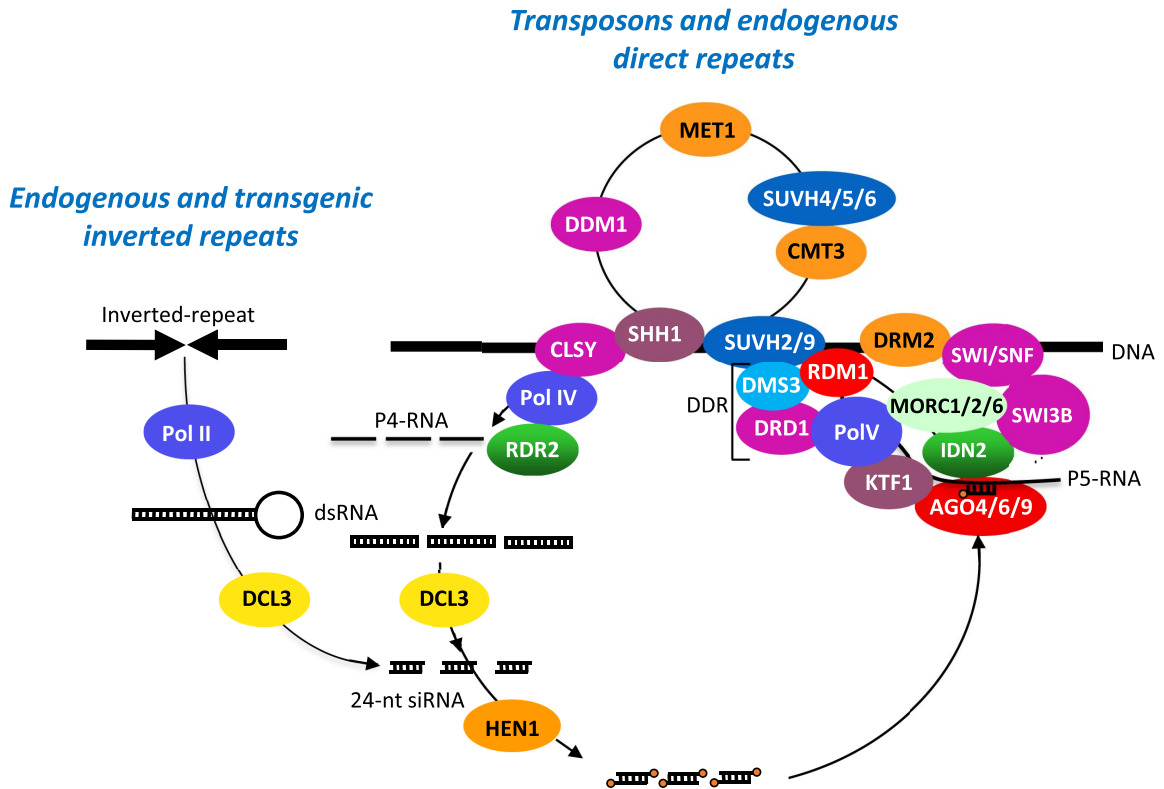
The discovery that 24-nt siRNAs mediate transgene TGS *in cis* and subsequently *in trans* allowed elucidating unexplained endogenous gene silencing events. For example, the promoter of the *FWA* gene contains two direct repeats that are kept methylated in vegetative tissues, causing *FWA* expression to be silenced. Additional copies of the *FWA* gene introduced into wild-type plants become immediately silenced, and their direct repeats are *de novo* methylated as if the silent endogenous *FWA* locus was acting *in trans* to silence homologous copies [21]. Using *FWA* transformation as a *de novo* trans-methylation assay revealed that *FWA* expression is regulated by RdDM [22]. The fact that the *FWA* promoter contains two direct repeats and not an inverted repeat implies that additional RdDM steps allowing the production of dsRNA from direct repeats exist upstream the processing of 24-nt siRNAs (Figure 2).

Elucidating the mechanism underlying transgene TGS also shed light on the paramutation

phenomenon originally discovered in maize in the 1950s [23–25]. Paramutation is the transfer of epigenetic information from a silent allele (referred to as paramutator) to an active homologous allele (referred to as paramutable), leading to the transformation of the paramutable allele into a paramutator allele with 100% efficiency, i.e. it is now capable of silencing paramutable alleles *in trans*. In the most studied locus, the maize *b1* locus, seven noncoding tandem repeats (TRs) are located ~100 kb upstream of the transcription start site of the *b1* gene in paramutagenic alleles. Maize lines lacking TRs do not undergo paramutation, while an allele carrying only three copies is paramutagenic but unstable and less penetrant [26]. The upstream region of *b1* is transcribed and produce siRNAs corresponding to the TRs. However, because these repeats are arranged in tandem and not as inverted repeats, similar to the situation with the Arabidopsis *FWA* gene, the production of siRNAs requires an RNA-dependent RNA polymerase (MOP1, the maize ortholog of Arabidopsis RDR2), which is essential for paramutation. Remarkably, a non-paramutagenic allele carrying a single copy also produce siRNAs, indicating that *b1* siRNAs participate to paramutation but are not sufficient. One possibility is that TR siRNAs can only establish a silenced state on particular chromatin conformation, and that such conformation requires the presence of repeats [23].

The mechanism of TGS *in trans* mediated by inverted-repeat transgenes producing dsRNA also explains endogenous gene silencing events occurring only in particular cultivars or in particular hybrids. For example, the *PAI* gene family is organized as three singlet genes in most Arabidopsis accession, including Col, whereas six other accessions, including Ws, exhibit an inverted-repeat in the promoter of the *PAI1* gene, likely resulting from a recent duplication, which promotes the production of dsRNA from the promoter region and the TGS of the *PAI2* and *PAI3* genes *in trans* [27, 28]. A similar situation is observed for the *FOLT* gene family, which arrangement and expression vary among Arabidopsis accessions, causing inter-accessions incompatibilities [29].

The presence of inverted repeats or tandem repeats may not be the only way to induce TGS at a given locus. Insertion of a transposable element (TE) nearby the promoter of a gene can cause the deregulation of this gene, and in some cases its silencing.



**Figure 2.** Model for TGS/RdDM. Silent and methylated endogenous sequences such as transposons or direct repeats bind the H3K9 methylation reader SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) and a chromatin remodeler of the CLSY family, allowing recruitment of Pol IV. The short ssRNAs made by Pol IV (P4-RNAs, 30–45-nt in length) are transformed into short dsRNA by RDR2, which physically interacts with Pol IV. These short dsRNAs are each processed into a 24-nt siRNA by DCL3. Endogenous sequences that do not undergo TGS/RdDM naturally can also be subjected to TGS/RdDM *in trans* by endogenous or transgenic inverted repeats. After methylation of the 24-nt RNA duplexes by HEN1, the guide strand is loaded into members of the AGO4-clade. AGO4 interacts with the its largest subunit of Pol V, which is recruited to methylated DNA by the catalytically inactive SUVH2/9 histone H3K9 methylases and the DDR complex (including the plant-specific protein RDM1, the solo hinge protein DMS3 involved in the structural maintenance of chromosome, and the chromatin remodeler DRD1). KTF1 secures the interaction between AGO4 and Pol V scaffold transcripts, allowing the pairing of AGO4-bound Pol IV-dependent 24-nt siRNAs with Pol V scaffold transcripts. The resulting dsRNA structure binds IDN2, which interacts with the MORC1/2/6 complex and a SWI/SNF complex involving the chromatin remodeler SWI3 to position nucleosomes, which contributes to condense chromatin, facilitate DNA methylation and repress transcription by Pol II. RDM1 likely helps AGO4 to recruit DRM2 to perform *de novo* DNA methylation. At CHH sites, the maintenance of methylation requires the continuous action of 24-nt siRNAs, while DNA methylation is maintained at CG and CHG sites by the MET1 DNA methyltransferase (CG sites) and the coordinated action of the CMT3 DNA methyltransferase and the catalytically active SUVH4/5/6 histone H3K9 methylases (CHG sites). Eventually, the chromatin remodeler DDM1 also contributes to maintaining DNA methylation.

In the case of the *WIP1* gene in melon, it was proposed that *WIP1* silencing is caused by the spreading of DNA methylation from a methylated TE of the *hAT* family inserted downstream of the *WIP1* gene [30]. Although this scenario is plausible, it is also possible that *WIP1* silencing is caused by transcription of the *WIP1* promoter originating in the nearby TE, leading to the production of dsRNA through the action of an RDR, and subsequent processing into siRNAs that provoke TGS and DNA methylation.

In the light of these findings, it may be interesting to re-examine the *peloria* variant of *Linaria vulgaris*. It was originally described as a pure epigenetic variant exhibiting reversible hypermethylation of the promoter region of the *Lcyc* gene [1]. However, the presence of TEs or repeats around the *Lcyc* gene was never investigated. Therefore, it remains possible that *peloria* is a true mutant carrying a rearranged *Lcyc* promoter or a TE inserted near the *Lcyc* gene, causing siRNA production and DNA methylation and TGS. If true, this scenario would explain how DNA methylation of the *Lcyc* promoter is maintained at CHH sites through generations. Indeed, whereas CG and CHG methylation are conservative and maintained by the MET1 DNA methyltransferase (CG) and the coordinated action of the CMT3 DNA methyltransferase and KYP Histone H3K9m2 methylase (CHG), CHH methylation requires the constant action of siRNAs to guide the DRM2 *de novo* DNA methyltransferase.

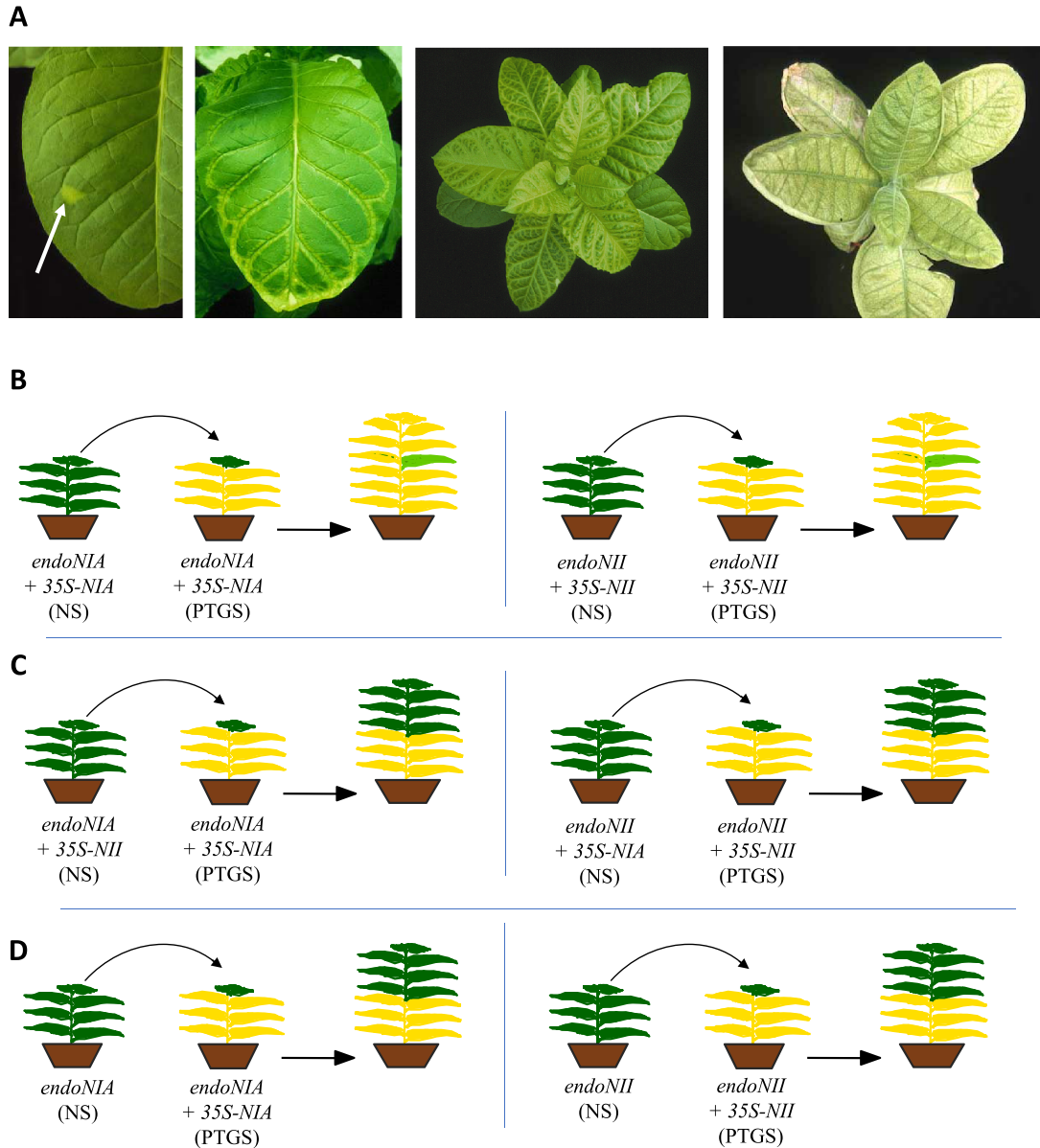
#### 4. The discovery of transgene PTGS

Two publications simultaneously revealed a silencing phenomenon that differed from those described so far [11, 14], and which turned out later to occur at the post-transcriptional level. In these reports, the authors showed that introduction of extra-copies of the *CHS-A* gene into *Petunia* can result in the silencing of both transgenic and endogenous copies, a phenomenon originally referred to as co-suppression [5, 6]. Similar results were obtained when introducing ectopic copies of other endogenous genes in various plant species (tobacco, tomato) [31–34], rapidly indicating that this phenomenon is not specific to *CHS* genes or *Petunia*. Remarkably, all reported cases of co-suppression involved transgenes expressed under the control of the strong viral 35S promoter. The silenced transgenes and homologous endogenous

genes remained correctly transcribed in the nucleus, but the corresponding mRNAs did not accumulate in the cytoplasm. Subsequently, co-suppression was referred to as PTGS. A direct correlation between the level of expression of the transgenic copy and the frequency of silencing was observed [35, 36] as if the cell was sensing that the amount of a specific RNA exceeded a threshold and decided to degrade them. Finally, degradation was confirmed when identifying siRNAs corresponding to the missing mRNAs [37].

PTGS was originally observed when introducing transgenes sharing homology with the transcribed sequence of endogenous genes, but it could also occur when introducing transgenes that do not show any homology with endogenous sequences [38]. This indicates that during co-suppression, transgenes act like inducers of PTGS while endogenous genes only behave as targets. Again, the intensity of PTGS of exogenous transgenes correlated with the level of transgene RNA, supporting the threshold hypothesis, but asking how does the cell counts specific RNAs, including molecules that do not exist naturally in the plant, and which type of RNAs are counted.

A striking feature of transgene PTGS is that it is dynamic and systemic, and erased during seed maturation. Indeed, dry seeds do not show PTGS, which starts after imbibition [39]. Once triggered, it lasts until the end of the plant's life, and plants transmit to their progeny their capacity to trigger PTGS, not their PTGS state. As a result, PTGS is triggered at each generation. Certain lines trigger PTGS very rapidly after germination, suggesting that it occurs synchronously in every cell. However, the use of lines that trigger PTGS lately revealed that PTGS starts locally, and then spreads through the vasculature to eventually invade systemically the entire plant [40] (Figure 3A). Grafting experiments involving non-silenced transgenic scions and silenced transgenic rootstocks revealed that a sequence-specific signal is transmitted from silenced tissues to non-silenced tissues (Figure 3B and C), explaining the natural spreading of transgene PTGS throughout the plant [41]. Such propagation was observed in cases of co-suppression (*p35S-NIA* and *p35S-NII* lines, see Figure 3) but also in lines carrying a *p35S-GUS* transgene that do not share any homology with the plant genome. The systemic aspect of PTGS was rapidly confirmed by



**Figure 3.** Transgene PTGS is systemic and sequence-specific. (A) Progressive establishment of transgene PTGS during somatic development. Co-suppression of tobacco *NIA* endogenous genes and transgenes is visualized by chlorosis. At each generation, the phenomenon starts as a spot on one leaf. Then, it progressively invades the entire plant by propagating through the vascular tissues. The same pattern is observed for every *NIA* transgenic line. Co-suppression of tobacco *NII* endogenous genes and transgenes occurs in the same manner. (B) Transgene PTGS is transmitted from silenced (PTGS) rootstocks to non-silenced (NS) scions carrying the same transgene (left: *p35S-NIA* plants, right: *p35S-NII* plants). (C) Systemic transgene PTGS is sequence-specific. Transgene PTGS is not transmitted from silenced *p35S-NII* rootstocks to non-silenced *p35S-NIA* scions (left) or from silenced *p35S-NIA* rootstocks to non-silenced *p35S-NII* scions (right). (D) Systemic transgene PTGS only occurs in the presence of a transgene. Wild-type scions carrying the endogenous *NIA* and *NII* genes but no *p35S*-driven homologous transgenes do not trigger PTGS when grafted onto silenced *p35S-NIA* (left) or *p35S-NII* (right) rootstocks.



introducing extra *p35S-GFP* transgene copies locally in a non-silenced *p35S-GFP* transgenic plant and observing the triggering of *GFP* PTGS in systemic tissues [42]. Remarkably, the use of co-suppression PTGS systems involving a transgene homologous to endogenous *NIA* or *NII* genes revealed that non-silenced transgenic scions can undergo PTGS when grafted onto silenced transgenic rootstocks, but not wildtype plants carrying only the endogenous genes, indicating that endogenous genes and transgenes have different competence for PTGS [43] (Figure 3D). Whether wildtype endogenous genes can undergo PTGS in the absence of homologous transgenes is discussed later.

To decipher the molecular mechanism underlying transgene PTGS, genetic screens were set up to identify mutants that either suppress or increase PTGS efficiency. Together, they define a model where transgene PTGS is divided in four steps: primary initiation, amplification, propagation and re-initiation, which makes it extremely efficient and systemic (Figure 4).

#### 4.1. Primary initiation

This step is still quite mysterious. Nevertheless, there is a consensus for transgene aberrant (ab)RNAs being at the root of the process. Firstly, because mutations in the RNA quality control (RQC) pathway, which degrades abRNAs, enhance PTGS [44–50], and secondly because uncapped and deadenylated RNAs were shown to be preferential substrates for cellular RDRs, in particular RDR6 [51–53]. Therefore, it is assumed that RDR6 transforms transgene abRNAs in dsRNA to initiate PTGS. Likely, this process occurs in a limited number of cells, either because of an unintended burst of transgene transcription, resulting in amounts of transgene abRNAs that exceed the degradation capacity of RQC, or because of an unintended failure of RQC functioning.

#### 4.2. Amplification

The cytoplasmic RNase type III enzymes DCL4 and DCL2 cut dsRNA into 21- and 22-nt siRNA duplexes, respectively [54, 55]. Those duplexes are methylated by HEN1 in their overhang 3' extremities to protect it from uridylation and degradation [56]. Then, siRNA duplexes are loaded into proteins from the ARGONAUTE (AGO) family, mostly AGO1 [57, 58].

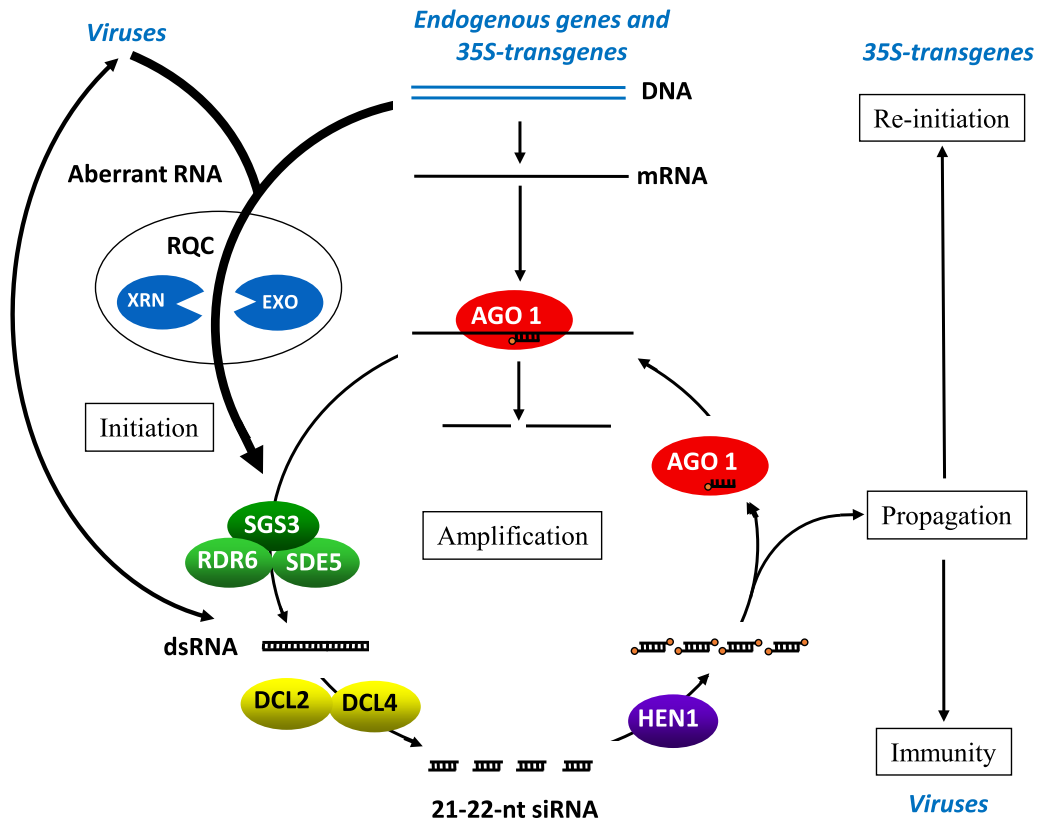
The passenger strand of the siRNA duplex is cleaved and eliminated, allowing the annealing of the guide siRNA strand with complementary transgene mRNA molecules and their cleavage owing to the RNaseH activity of AGO proteins. After cleavage by AGO/siRNA complexes, transgene RNA fragments are degraded by exonucleases. However, transgene mRNA bound to an AGO1/22-nt siRNA complex can be transformed into dsRNA by RDR6, leading to the production of secondary siRNAs [59, 60] and their subsequent loading into AGO proteins, which maximizes the elimination of transgene mRNA.

#### 4.3. Propagation

PTGS eventually becomes systemic owing to a mobile RNA signal that can move from cell-to-cell and at long distance through the vascular tissues. Although the systemic nature of PTGS was discovered 25 years ago [41, 42], little is still known about the propagation step. In fact, there is still debate whether the RNA signal is dsRNA or siRNA. The movement of siRNAs was clearly demonstrated [61, 62], but the possibility that the PTGS signal is made of dsRNA has not been formally ruled out. In addition, it is not known if the RNA signal moves alone or associated with cargo proteins, and what regulates this movement. So far, mutants impaired in the PTGS propagation step have not been identified, likely because the elements involved in the movement of the PTGS signal are essential for the plant.

#### 4.4. Reinitiation

PTGS becomes systemic not only because of the existence of a mobile RNA signal but also because transgenes are capable of reinitiating PTGS in cells where it has not been initiated by RQC saturation or deficiency. This step also is poorly understood, mostly because it requires performing reciprocal grafting experiments (wild-type recipient tissues grafted onto mutant emitting tissues vs mutant recipient tissues grafted onto wildtype emitting tissues) to determine at which step a given mutation impairs systemic PTGS. As expected, mutations in AGO1, DCL2/DCL4, RDR6 and SGS3 in either emitting or recipient tissues impair PTGS because these are core PTGS proteins required in every cell to execute PTGS [63–66]. So far,



**Figure 4.** Model for PTGS. Aberrant RNAs are considered at the root of PTGS. They are naturally produced by endogenous genes, transgenes and viruses, and should be eliminated by the RNA quality control (RQC) pathways. However, when their amounts exceed the capacity of RQC or when RQC becomes accidentally dysfunctional in a cell, the aberrant RNAs are converted to dsRNA by RDR6 to initiate PTGS. In the case of RNA viruses, intermediates of viral replication consist in dsRNA that could potentially initiate PTGS too. The cytoplasmic RNase type III enzymes DCL4 and DCL2 cut dsRNA into 21- and 22-nt siRNA duplexes, which are methylated by HEN1. Complementary mRNAs are either cleaved by AGO1/siRNA complexes or copied by RDR6 assisted with the RNA-binding protein SGS3 and the RNA export protein SDE5 to produce more dsRNA, leading to the production of secondary siRNAs, thus maximizing mRNA elimination. PTGS eventually spreads owing to a mobile RNA signal that can move from cell-to-cell and at long distance through the vascular tissues. Likely, the RNA signal is made of siRNAs, which mobility was clearly demonstrated. In the case of viruses, if siRNAs spread faster than the virus itself, they confer immunity to the non-infected tissues. In the case of transgenes driven by the 35S promoter, PTGS spreads to the entire plant because transgenes are capable of reinitiating PTGS in cells where it has not been initiated by RQC saturation or deficiency. This capacity requires the histone H3K4me3 demethylase JMJ14 and the NAC-domain transcription factor NAC52, which protect transgene sequences from *de novo* DNA methylation. Endogenous sequences are not protected sequences from *de novo* DNA methylation by JMJ14 and NAC52, and are not capable of undergoing systemic PTGS, which likely explain how plants survive to local and occasional RQC deficiency.

no mutations affecting specifically the capacity of the emitting tissues to transmit PTGS to recipient tissues have been identified. In contrast, several mutations affecting specifically the capacity of the recipient tissues to trigger PTGS when grafted onto emitting tissues have been identified. However, the role of the corresponding genes remains puzzling. It comprises genes encoding part of the TGS machinery: *PolIV*, *RDR2*, *DCL3*, *AGO4* [63,65], and two genes, *JMJ14* and *NAC52*, encoding a histone H3K4me3 demethylase and a NAC-domain transcription factor. *JMJ14* and *NAC52* are not required for PTGS execution nor for the production of the RNA silencing signal, but are required for transgene PTGS to become systemic [64,67]. Remarkably, DNA methylation increases at transgenic sequences in the absence of *JMJ14* or *NAC52*, but not at any endogenous sequences [64]. Likely, it is not the absence of *JMJ14* or *NAC52* but transgene hypermethylation that prevents the reinitiation of PTGS. Indeed, PTGS reinitiation occurs normally in a *jmj14* mutant background when hypermethylation is prevented by mutating simultaneously the *DRM2* and *CMT3* DNA methylases [64]. How *JMJ14* specifically attenuates transgene DNA methylation and promotes systemic transgene PTGS is still not understood.

## 5. Natural PTGS phenomena

The mechanism at work when plants recover from virus infection is very similar to that involved in transgene PTGS (Figure 4). Likely, the enormous amount of viral RNAs produced in the first infected cells includes sufficient abRNAs to saturate RQC, leading to the production of primary dsRNA that are processed in the cytoplasm by *DCL4* and *DCL2*, resulting in 21- and 22-nt primary siRNA duplexes that are loaded onto proteins from the ARGONAUTE (AGO) family, mostly *AGO1*, but also *AGO2*, *AGO5*, *AGO7* and *AGO10* [68]. Alternatively, or concomitantly, the viral dsRNA that is formed during virus replication could be directly used by DCLs to produce the primary siRNAs. Once siRNAs are loaded into AGOs, they guide the cleavage of viral single-stranded RNAs. Similar to what happens during transgene PTGS, the *AGO1/22-nt* siRNA complex specifically promotes the production of additional dsRNA by *RDR6*, leading to the production of secondary siRNAs and their subsequent loading onto AGO proteins,

which maximizes the elimination of viral RNAs from the plant cell.

Whereas transgenes driven by the viral 35S promoter are often silenced by PTGS, examples of transgenes driven by endogenous promoters and undergoing PTGS are missing. Nevertheless, a handful of protein-coding genes (PCGs), in certain varieties, have been shown to produce siRNAs at levels that allow degrading mRNAs by PTGS. Remarkably, these varieties exhibit genomic rearrangements involving duplication events, which allow producing dsRNA, whereas regular varieties exhibit a single gene copy that does not produce siRNAs [7, 69–72]. It is likely that such rearrangements are tolerated because they affect dispensable genes. This implies that (i) during evolution, endogenous protein-coding genes have been shaped to avoid producing excessive amounts of siRNAs, (ii) cells undergoing duplication events that provoke the production of siRNAs from essential genes do not survive. Importantly, PTGS events involving transgenes always are systemic, whereas the few known examples of PCGs undergoing PTGS only show localized silencing, confirming what was observed when grafting wildtype plants on transgenic plants silenced by PTGS [43].

The results described above suggested that only genomic rearrangements leading to the production of dsRNA could promote the production of siRNAs from PCGs. However, most RQC-deficient mutants have long been known to die at a very early stage of development, suggesting that the abRNAs produced by PCGs could enter the PTGS pathway when they are not eliminated by RQC, thus producing siRNAs that could target functional homologous mRNA for destruction, which eventually kills the plant. Supporting this hypothesis, the lethality of RQC-deficient mutants was rescued by mutating the PTGS pathway simultaneously [50, 73]. The inventory of siRNAs in *Arabidopsis* mutants impaired in decapping or exoribonuclease activities revealed that thousands of PCGs produce siRNAs (coined *rqc*-siRNAs) when plants are deficient for RQC. Because RQC null alleles are embryolethal or die at a very early stage of development, a complete view of PCGs capable of producing siRNAs when RQC is impaired cannot be obtained using constitutive mutants. Analysis of mutants that partially impair RQC extended the list of siRNA-producing PCGs [50, 73–75]. Moreover, another category of siRNAs derived from PCGs was

identified when plants are infected by viruses [76]. These siRNAs were called virus-activated siRNAs (va-siRNAs). There is a large overlap between rqc-siRNAs and va-siRNAs, suggesting that viruses could provoke the production of siRNAs from PCGs by expressing proteins that inhibit RQC or by producing abRNAs up to a level that saturates RQC pathway and trigger the entry of PCG-derived abRNAs into the PTGS pathway. Overall, ~5000 PCGs among the ~27,000 *Arabidopsis* PCGs can produce siRNAs. Up to now, it is not known if only these PCGs are capable of producing siRNAs or if their number is underestimated due to the restricted analysis of tissues surviving in RQC-deficient mutant or virus infected plants.

It is possible that, during evolution, RQC was adjusted to the amount of aberrant RNAs produced by endogenous PCGs so that PCGs do not produce siRNAs in RQC-proficient cells. Alternatively, the local or transient impairment of RQC may not have any dramatic consequences on the entire plant because, unlike transgenes, endogenous PCGs are incapable of undergoing systemic PTGS. This may have to do with the distinct epigenetic regulation of transgenes and endogenous PCGs. Indeed, whereas DNA methylation of endogenous PCGs is limited by the histone H3K9me2 demethylase IBM1, DNA methylation of transgenes driven by the 35S promoter is limited by the histone H3K4me3 demethylase JMJ14 but not IBM1. JMJ14 is not required for PTGS execution nor for the production of the RNA silencing signal, but is required for transgene PTGS to become systemic [64]. How JMJ14 specifically attenuates transgene DNA methylation and promotes systemic transgene PTGS is still not understood. Nevertheless, the distinct epigenetic regulations of transgenes and endogenous PCGs implies that a local production of endogenous siRNAs has no consequence on endogenous PCG expression in the rest of the plant, whereas a local production of transgene siRNAs activates systemic transgene PTGS. In the current PTGS model, once siRNAs are produced in certain cells, they move to other cells but only transgenes, not endogenous PCGs, can reinitiate PTGS in cells receiving the siRNAs because of the specific regulation of transgenes by JMJ14. As a result, transgenes driven by the 35S promoter are prone to produce 21- and 22-nt siRNAs and trigger systemic PTGS, whereas endogenous PCGs producing 21- and 22-nt siRNAs and

exhibiting signs of PTGS are scarce and only show localized PTGS [7, 69–72].

## 6. Natural phenomena involving both TGS and PTGS

In the above sections, TGS and PTGS are defined as independent phenomena, acting on different targets. Although this is generally the case, transposable elements (TEs) set apart, involving TGS and PTGS at the different steps of their epigenetic control. The maintenance of TE silencing involves the RdDM mechanism described in Figure 2. However, despite this active maintenance mechanism [18], TEs can sometime be reactivated, for example in pollen where DDM1 is not expressed [77], or in mutants where DDM1 or MET1 are artificially impaired [78, 79]. In these cases, TEs are subjected to PTGS first, before being re-silenced by TGS. Briefly, TE transcripts either undergo ribosome stalling, which causes RNA cleavage [80], or are cleaved by specific miRNAs, allowing cleavage products to be transformed into dsRNA by RDR6 [81]. These TE-derived dsRNA are processed by DCL4 and DCL2 into 21- and 22-nt siRNAs, which are loaded into AGO1 and AGO2. However, this is insufficient to suppress TE expression. Eventually, DCL3-mediated processing of TE-derived dsRNA into 24-nt siRNAs that are loaded into AGO4-clade AGOs allows guiding DNA methylation to TE reactivated copies, resulting in the re-initiation of TGS [82], which can subsequently be maintained by the canonical RdDM pathway (Figure 2).

## 7. Concluding remarks

Genomes are constantly evolving. At each generation, crossing-overs, which are mandatory for meiosis to occur correctly [83], juxtapose sequences that were not adjacent in the parental genomes. Crosses between different plants assemble in their progeny two genomes that can strongly differ in their sets of long and small non-coding RNAs and associated epigenetic marks [84, 85]. When subjected to genotoxic stresses, genomes can be damaged, and the following repair can be perfect or imperfect. Genomes can also undergo partial or complete duplication [86]. TEs can move or duplicate, and the

new inserted copies can disrupt a gene or insert nearby and modify its expression [87]. Finally, some pathogens can insert part of their genetic material in the plant genome, and have consequences similar to TE insertions. For all these reasons, distinguishing self from modified/damaged self and non-self is essential. The DNA sequence is clearly insufficient to provide such information. Likely, the epigenetic marks that decorate DNA allow this identification process. Better, it could simply be the absence of native epigenetic marks that allows identifying non-self or modified/damaged self. Indeed, new DNA sequences (transgene DNA, pathogen DNA, new transposable element copy, damaged/repared DNA, duplicated genomic fragment) generally are naked. To form chromatin, DNA assembles with histones, but how are chosen the epigenetic marks that decorate the histones that associate with new DNA? By default, histones bearing particular decorations could associate with naked DNA, marking these new sequences as potentially dangerous, requiring immediate silencing, or at least surveillance for a long period of time before they are considered as safe and thus self.

### *Version française*

#### **1. Préhistoire de l'extinction de gènes**

Le terme d'extinction de gènes a été inventé lorsqu'on a réalisé que les plantes transgéniques n'exprimaient pas toujours les transgènes introduits. Ces phénomènes d'extinction de gènes ne sont pas apparus comme une conséquence de la transgénèse. Ils existaient déjà naturellement, mais n'étaient pas connus en tant que tels. Il serait trop long d'énumérer tous les phénomènes de ce type recensés avant l'ère de la transgénèse, aussi cette section introductive se concentrera sur trois cas représentatifs de la manière dont les gènes endogènes peuvent être réduits au silence au niveau transcriptionnel ou post-transcriptionnel, et la façon dont les plantes éliminent le matériel génétique d'éléments exogènes (Figure 1).

La plus ancienne description d'une variabilité phénotypique instable est due à Linée qui a décrit en 1749 le variant *peloria* de la plante *Linaria vulgaris*. Alors que la plante de référence montre des fleurs à symétrie bilatérale, le variant *peloria* montre des fleurs à symétrie radiale (Figure 1A).

#### **Conflicts of interest**

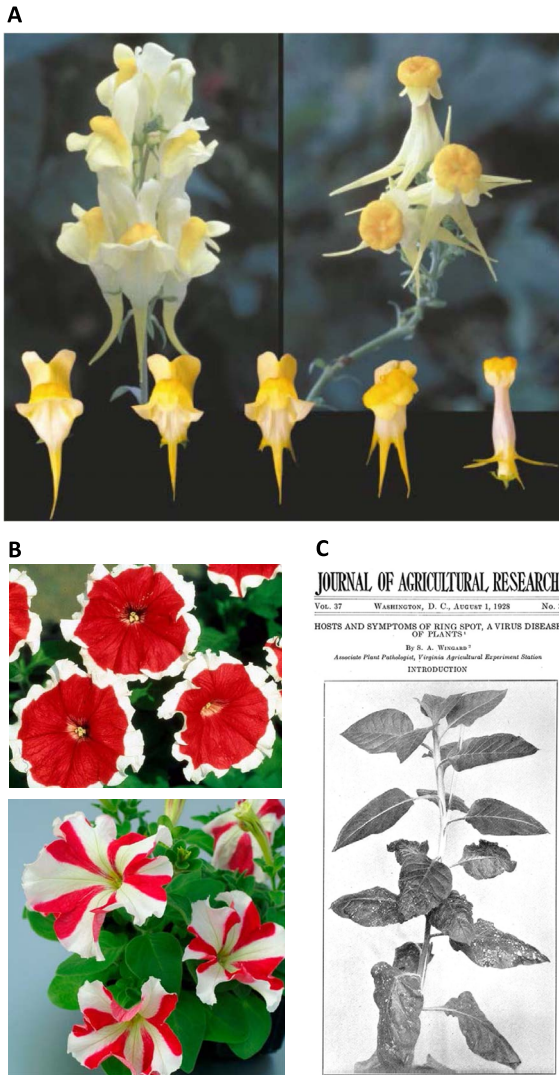
The author has no conflict of interest to declare.

#### **Acknowledgements**

First, I would like to thank Michel Caboche for welcoming me in his lab in 1985 and transmitting me his passion for plant genetics. His mentoring was essential to my career. I also thank Jean-Pierre Bourgin, Francine Casse, Yves Chupeau, Georges Pelletier and Pierre Rouzé for their support and encouragements at a time gene silencing was still a mystery. Thanks to Jean-Luc Rossignol and Godeleine Faugeron for sharing their ideas on epigenetics when most people did not know what it was. A great thank to the members of my group for enduring my interest for aberrant segregations, and in particular Taline Elmayan for 30 years of faithful and enthusiastic collaboration. Finally, I would like to thank the members of the international gene silencing community, and in particular Marjori and Antonius Matzke for fruitful discussions around the first (silenced) transgenic plants.

Occasionnellement, le variant réverte vers un phénotype sauvage pendant le développement somatique, ce qui en fait le premier variant épigénétique jamais décrit. En effet, la modification florale n'est pas due à une mutation mais à une modification épigénétique (épimutation). Dans le cas du variant *peloria*, la modification consiste en une hyperméthylation de l'ADN, qui provoque une extinction transcriptionnelle (en anglais : *transcriptional gene silencing* ou TGS) du gène endogène *Lycy* [1]. Comme toute épimutation, l'épimutation *peloria* n'existe pas dans un état unique. En effet, il existe une gamme de phénotypes intermédiaires, dont l'intensité est corrélée au degré de méthylation de l'ADN et d'expression résiduelle du gène endogène *Lycy* (Figure 1A).

La plus ancienne description d'un phénotype panaché date de 1896 et concerne des cultivars horticoles de *Petunia hybrida* présentant des fleurs bicolores. *Petunia hybrida* est un hybride inter-spécifique entre *Petunia axillaris*, qui présente des fleurs blanches, et *Petunia integrifolia*, qui présente des fleurs rouges/pourpres [2,3]. Parmi les hybrides obtenus, le cultivar *Picotée* présentait des fleurs avec des



**FIGURE 1.** Exemples représentatifs de phénomènes d'extinction observés dans la nature. (A) Fleurs du type sauvage *Linaria vulgaris* (en haut à gauche) et du variant épigénétique *peloria* (en haut à droite). En dessous sont représentés les phénotypes intermédiaires observés au cours du développement somatique, allant du type sauvage (à gauche) au phénotype *peloria* le plus fort (à droite). L'intensité du phénotype *peloria* est corrélée au degré de méthylation et d'extinction du gène endogène *Lcyc* par TGS. (B) Cultivars horticoles de *Petunia hybrida* présentant une extinction des gènes endogènes *CHS-A* par PTGS, résultant en des fleurs bicolores.

**FIGURE 1. (cont.)** En haut : le cultivar *Picotée* présente des fleurs à centre rouge et à bord blanc. En bas : le cultivar *Star* présente des secteurs rouges et blancs alternant dans un motif radial. (C) Plantes de tabac se remettant d'une infection virale grâce au PTGS. Au bas de la plante, les feuilles les plus anciennes présentent de graves symptômes d'infection par un virus appelé « tobacco ringspot virus ». Les symptômes disparaissent progressivement au fur et à mesure que la plante grandit jusqu'à ce qu'elle produise des feuilles supérieures exemptes de virus et immunisées contre une surinfection par le même virus.

centres rouges et des marges blanches, tandis que le cultivar *Star* présentait des secteurs rouges et blancs alternant dans un motif radial [4] (Figure 1B). Ces motifs ont été reproduits par transgénèse en introduisant des copies supplémentaires du gène *CHS-A*, ce qui a conduit à l'extinction post-transcriptionnelle des copies endogènes et transgéniques de *CHS-A* (en anglais : *post-transcriptional gene silencing* ou PTGS) [5, 6]. Plus récemment, il a été montré que le phénotype des cultivars *Picotée* et *Star* résultait de duplications endogènes du gène *CHS-A*, alors que les autres variétés de pétunias (de couleur uniforme) portent une seule copie du gène *CHS-A* [7].

La description la plus ancienne d'un phénomène d'extinction de gène affectant un élément exogène est due à Wingard qui a décrit en 1928 comment des plantes de tabac pouvaient se remettre d'une infection virale et devenir immunes à la surinfection par le même virus [8] (Figure 1C). Des décennies ont été nécessaires pour comprendre que le mécanisme de PTGS mis en œuvre pour éteindre les copies ectopiques de transgènes était également utilisé pour produire des petits ARN qui ciblent les longs ARN viraux pour les détruire [9].

## 2. Découverte du TGS des transgènes

Au début des années 80, les biologistes des plantes ont utilisé la capacité de la bactérie du sol *Agrobacterium tumefaciens* à transférer une partie de son ADN (appelé ADN-T) pour mettre au point une méthode universelle de transformation des plantes. La première plante génétiquement modifiée était

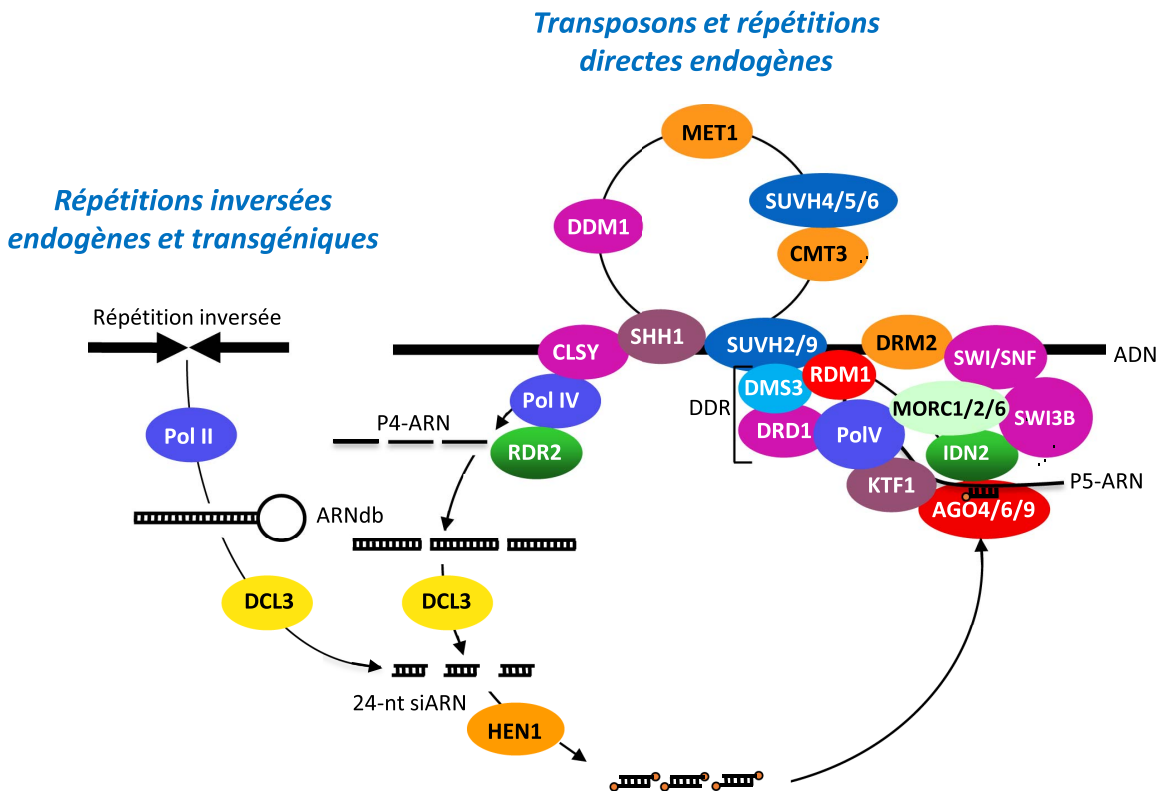
une plante de tabac portant un ADN-T modifié d'*Agrobacterium* dans lequel étaient insérés un gène bactérien de *néomycine phosphotransférase* et un gène d'*alcool déshydrogénase* de levure. Une plante transgénique a été régénérée, qui a transmis l'ADN-T à sa descendance [10]. Cependant, le gène de la *nopaline synthase* d'*Agrobacterium* porté par l'ADN-T n'était exprimé que dans 24% de la descendance (alors qu'on s'attendait à ce qu'il soit exprimé dans au moins 75% de la descendance), indiquant qu'une partie de la descendance portait l'ADN-T mais n'exprimait pas un de ses gènes. Les auteurs ont trouvé ce résultat déroutant et inattendu, mais n'y ont pas donné suite. Rétrospectivement, il semblerait que ce soit le premier exemple d'extinction de gène affectant une plante transgénique rapporté dans la littérature (Antonius et Marjori Matzke, communication personnelle), et il est remarquable qu'il se soit produit dans la première plante transgénique jamais obtenue.

Au cours des années suivantes, un nombre croissant de publications ont décrit l'obtention de plantes transgéniques qui exprimaient les transgènes introduits, mais aucune publication ne faisait état de problèmes d'expression, soit parce qu'aucune attention n'était portée aux plantes de la descendance qui n'exprimaient plus les transgènes, soit parce que les auteurs censuraient ce qui pouvait être considéré comme défavorable à la publication de leurs travaux. Une étape importante a été franchie lorsqu'on a observé que la transformation séquentielle de plantes de tabac par un second ADN-T pouvait conduire à l'extinction réversible des transgènes portés par le premier ADN-T et qui partageaient une homologie dans leurs séquences promotrices. En effet, ce phénomène d'extinction pouvait être réverté lors de la ségrégation du second ADN-T [11]. Il est important de noter que cette première étude a permis d'établir une corrélation entre la méthylation de l'ADN des séquences promotrices et une forme d'extinction qui s'est avérée par la suite être transcriptionnelle (TGS) (12). Étant donné que l'extinction des transgènes du premier ADN-T ne se produisait qu'en présence du second ADN-T, on a d'abord pensé que ce phénomène dépendait directement du nombre de copies, tout comme le phénomène MIP décrit chez le champignon *Ascobolus immersus* [13]. L'étude de plantes de pétunia transgéniques portant le gène de la dihydrofolate réductase *A1* du maïs a confirmé la

corrélation positive entre le nombre de copies et l'extinction et la méthylation [14]. De plus, cette étude a révélé qu'un locus de transgène *A1* qui subit spontanément le TGS en *cis* peut déclencher en *trans* le TGS d'un locus de transgène *A1* actif [14], expliquant ainsi comment le TGS peut se mettre en place lors de transformations séquentielles [11]. L'étude d'une plante de tabac portant de multiples copies méthylées et éteintes d'un plasmide portant des transgènes exprimés sous le contrôle des promoteurs viraux 19S et 35S a révélé que le TGS se manifestait en *trans* indépendamment de la position des transgènes cibles dans le génome et nécessitait uniquement une homologie du promoteur mais pas de la séquence codante [15]. Finalement, il a été montré que les loci transgéniques capables de subir le TGS spontanément en *cis* et de le transmettre ensuite en *trans* résultaient d'événements d'insertion générant une structure répétée inversée permettant la production d'ARN double brin (ARNdb) à partir des séquences promotrices du transgène [16, 17]. Ces ARNdb sont transformés en une classe de petits ARN interférents (siARN) de 24 nucléotides (nt), qui guident la méthylation de l'ADN *de novo* et le remodelage de la chromatine au niveau des loci homologues pour les éteindre en *trans*, un processus également appelé méthylation de l'ADN dirigée par l'ARN (en anglais : *RNA-directed DNA methylation* ou RdDM) [18]. Divers cribles génétiques ont été mis en place en utilisant des systèmes transgéniques à deux composants, permettant de disséquer le mécanisme sous-jacent au RdDM [18–20]. Un schéma de ce processus est représenté sur la Figure 2.

### 3. Phénomènes naturels de TGS

La mise en évidence du rôle des siARN de 24 nt dans le TGS des transgènes en *cis* et en *trans* a permis d'élucider des événements inexpliqués d'extinction de gènes endogènes. Par exemple, le promoteur du gène *FWA* contient deux répétitions directes qui sont maintenues méthylées dans les tissus végétatifs, ce qui entraîne l'extinction du gène *FWA*. Des copies supplémentaires du gène *FWA* introduites dans des plantes de type sauvage sont immédiatement éteintes, et leurs répétitions directes sont méthylées *de novo*, indiquant que le locus *FWA* endogène agit en *trans* pour éteindre les copies homologues [21]. L'utilisation de la transformation par *FWA* comme test



**FIGURE 2.** Modèle de TGS/RdDM. Les séquences endogènes silencieuses et méthylées telles que les transposons ou les répétitions directes se lient au lecteur de méthylation H3K9 SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) et à un remodelleur de chromatine de la famille CLSY, permettant le recrutement de Pol IV. Les courts ARNs fabriqués par la Pol IV (P4-RNAs, 30–45-nt de longueur) sont transformés en courts ARNdb par RDR2, qui interagit physiquement avec la Pol IV. Ces courts ARNdb sont chacun transformés en un siARN de 24 nt par DCL3. Les séquences endogènes qui ne subissent pas naturellement la TGS/RdDM peuvent également être soumises à la TGS/RdDM en trans par des répétitions inversées endogènes ou transgéniques. Après méthylation des duplex d'ARN de 24 nt par HEN1, le brin guide est chargé sur des protéines AGO de la famille AGO4. AGO4 interagit avec la plus grande sous-unité de Pol V, qui est recrutée sur l'ADN méthylé par les méthylases d'histone H3K9 SUVH2/9 catalytiquement inactives et le complexe DDR (comprenant la protéine RDM1 spécifique aux plantes, la protéine charnière DMS3 impliquée dans le maintien structurel du chromosome, et le remodelleur de chromatine DRD1). KTF1 sécurise l'interaction entre AGO4 et les transcrits Pol V, ce qui permet l'appariement des siARN de 24 nt dépendants de la Pol IV liés à AGO4 avec les transcrits Pol V. La structure ARNdb qui en résulte se lie à IDN2, qui interagit avec le complexe MORC1/2/6 et un complexe SWI/SNF impliquant le remodelleur de chromatine SWI3 pour positionner les nucléosomes, ce qui contribue à condenser la chromatine, à faciliter la méthylation de l'ADN et à réprimer la transcription par Pol II. RDM1 aide probablement AGO4 à recruter DRM2 pour effectuer la méthylation de l'ADN de novo. Au niveau des sites CHH, le maintien de la méthylation nécessite l'action continue de siRNA de 24 nt, tandis que la méthylation de l'ADN est maintenue au niveau des sites CG et CHG par l'ADN méthyltransférase MET1 (sites CG) et l'action coordonnée de l'ADN méthyltransférase CMT3 et des histone H3K9 méthylases SUVH4/5/6 catalytiquement actives (sites CHG). Enfin, le remodelage de la chromatine DDM1 contribue également au maintien de la méthylation de l'ADN.



de *trans*-méthylation *de novo* a révélé que l'expression de *FWA* est régulée par RdDM [22]. Le fait que le promoteur de *FWA* contienne deux répétitions directes et non une répétition inversée implique que des étapes supplémentaires de RdDM permettant la production d'ARNdb à partir de répétitions directes existent en amont du traitement des siARN de 24 nt (Figure 2).

L'élucidation du mécanisme qui sous-tend le TGS des transgènes a également permis de faire la lumière sur le phénomène de paramutation découvert à l'origine chez le maïs dans les années 1950 [23–25]. La paramutation est le transfert d'information épigénétique d'un allèle éteint (appelé paramutateur) à un allèle homologue actif (appelé paramutable), conduisant à la transformation de l'allèle paramutable en allèle paramutateur avec une efficacité de 100%, l'allèle paramutateur devenant alors capable d'éteindre les allèles paramutables en *trans*. Dans le cas du locus le plus étudié, le locus *b1* du maïs, une séquence non codante répétée sept fois en tandem (en anglais *tandem repeats* ou TRs) est située ~100 kb en amont du site d'initiation de la transcription du gène *b1* dans les allèles paramutateurs. Les lignées de maïs dépourvues de TRs ne subissent pas de paramutation, tandis qu'une lignée portant seulement trois copies demeure paramutable mais est instable et moins pénétrante [26]. La région en amont de *b1* est transcrite et produit des siARN correspondant aux TRs. Cependant, parce que ces répétitions sont disposées en tandem comme c'est le cas pour le gène *FWA* d'Arabidopsis, la production de siARN nécessite une ARN polymérase ARN-dépendante (MOP1, l'orthologue du maïs de la protéine RDR2 d'Arabidopsis), qui est essentielle pour la paramutation. De façon remarquable, un allèle non paramutable portant une seule copie produit également des siARN, indiquant que les siARN du gène *b1* participent à la paramutation mais ne sont pas suffisants. Il est donc envisageable que les siARN ne puissent établir le TGS que sur une conformation particulière de la chromatine, et que cette conformation nécessite la présence de répétitions [23].

Le mécanisme de TGS en *trans* médié par des transgènes à répétition inversée produisant des ARNdb explique également des événements d'extinction de gènes endogènes se produisant uniquement dans des cultivars particuliers ou dans des hybrides particuliers. Par exemple, la famille des

gènes *PAI* comprend trois gènes sans répétitions dans la plupart des accessions d'Arabidopsis, alors que certaines accessions présentent une répétition dans le promoteur du gène *PAI1*, résultant probablement d'une duplication récente, qui provoque la production d'ARNdb à partir de la région du promoteur et le TGS des gènes *PAI2* et *PAI3* en *trans* [27, 28]. Une situation similaire est observée pour la famille de gènes *FOLT*, dont l'agencement et l'expression varient entre les accessions d'Arabidopsis, provoquant des incompatibilités inter-accessions [29].

La présence de répétitions inversées ou de répétitions en tandem n'est peut-être pas le seul moyen d'induire le TGS à un locus donné. L'insertion d'un élément transposable à proximité du promoteur d'un gène peut provoquer la dérégulation de ce gène, et dans certains cas son extinction. Dans le cas du gène *WIP1* du melon, il a été proposé que l'extinction de *WIP1* soit causée par la propagation de la méthylation de l'ADN à partir d'un transposon méthylé de la famille *hAT* inséré en aval du gène *WIP1* [30]. Bien que ce scénario soit plausible, il est également possible que l'extinction de *WIP1* soit causée par la transcription du promoteur de *WIP1* à partir du transposon voisin, conduisant à la production d'ARNdb par l'action d'une RDR, puis à la transformation en siARN qui provoque le TGS et la méthylation de l'ADN.

À la lumière de ces résultats, il serait intéressant de réexaminer le variant *peloria* de *Linaria vulgaris*. Celui-ci a été décrit à l'origine comme un pur variant épigénétique présentant une hyperméthylation réversible de la région promotrice du gène *Lcyc* [1]. Cependant, la présence de transposons ou de répétitions autour du gène *Lcyc* n'a jamais été étudiée. Par conséquent, il reste possible que *peloria* soit un véritable mutant portant un promoteur *Lcyc* réarrangé ou un transposon inséré près du gène *Lcyc*, provoquant la production de siARN et la méthylation de l'ADN et le TGS. Un tel scénario expliquerait comment la méthylation de l'ADN du promoteur *Lcyc* est maintenue sur les sites CHH au fil des générations. En effet, alors que la méthylation CG et CHG est conservée et maintenue par l'ADN méthyltransférase MET1 (CG) et l'action coordonnée de l'ADN méthyltransférase CMT3 et de l'Histone H3K9m2 méthylase KYP (CHG), la méthylation CHH nécessite l'action constante des siARN pour guider l'ADN méthyltransférase *de novo* DRM2.

#### 4. Découverte du PTGS des transgènes

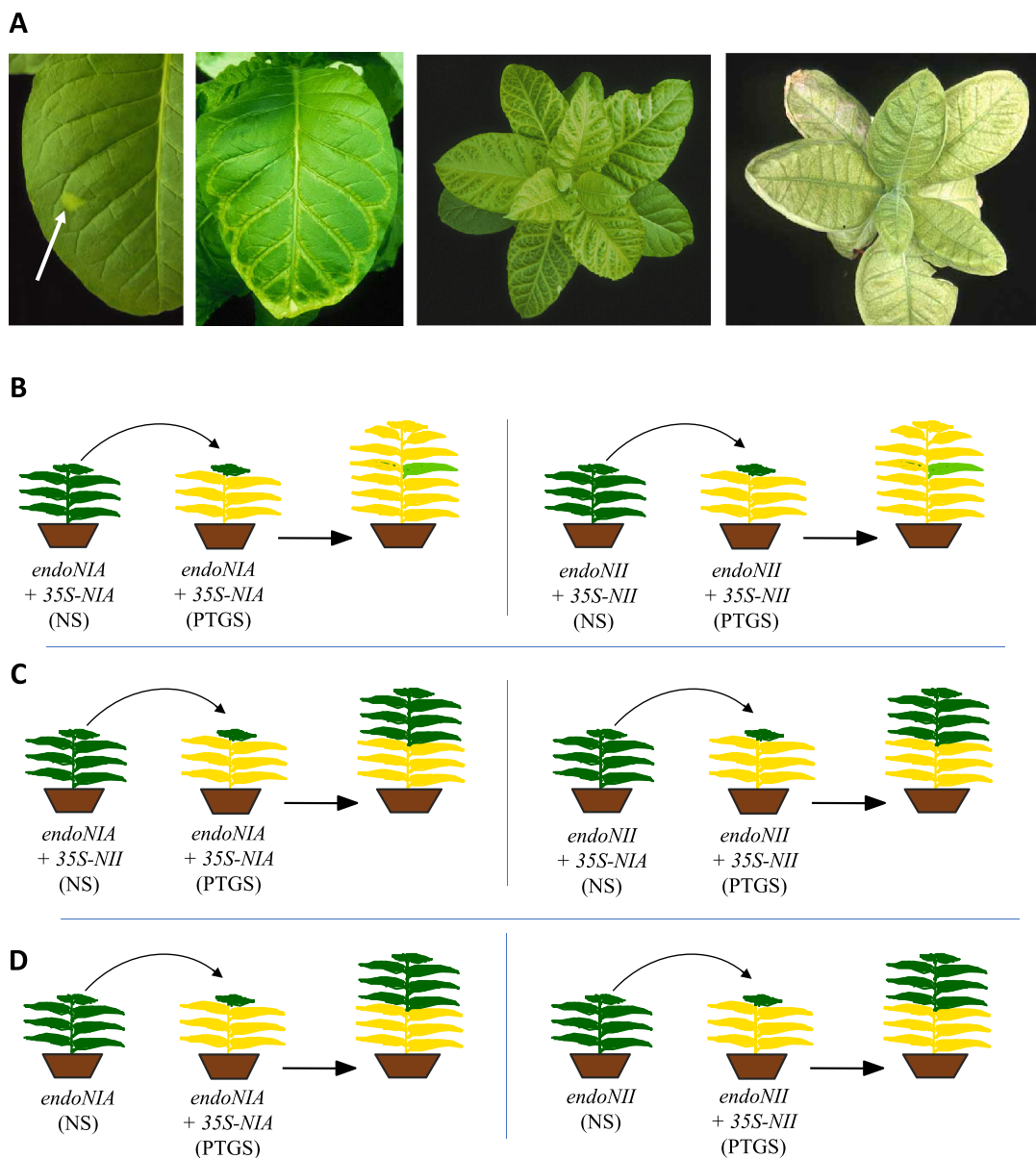
Deux publications ont révélé simultanément un phénomène d'extinction différent de ceux décrits jusqu'à présent [11, 14], et qui s'est avéré par la suite se produire au niveau post-transcriptionnel. Dans ces publications, les auteurs ont montré que l'introduction de copies supplémentaires du gène *CHS-A* dans le pétunia peut entraîner l'extinction des copies transgéniques et endogènes, un phénomène initialement appelé co-suppression [5, 6]. Des résultats similaires ont été obtenus lors de l'introduction de copies ectopiques d'autres gènes endogènes dans diverses espèces végétales (tabac, tomate) [31–34], indiquant rapidement que ce phénomène n'était pas spécifique aux gènes *CHS* ou au pétunia. De façon remarquable, tous les cas de co-suppression rapportés impliquaient des transgènes exprimés sous le contrôle du puissant promoteur viral 35S. Les transgènes éteints et les gènes endogènes homologues restaient correctement transcrits dans le noyau, mais les ARNm correspondants n'étaient plus accumulés dans le cytoplasme. Par la suite, la co-suppression a été désignée sous le nom de PTGS. Une corrélation directe entre le niveau d'expression de la copie transgénique et la fréquence de l'extinction a été observée [35, 36] comme si la cellule détectait que la quantité d'un ARN spécifique dépassait un seuil et décidait de les dégrader. Enfin, la dégradation a été confirmée lors de l'identification des siARN de 21 et 22 nt correspondant aux ARNm manquants [37].

Le PTGS a été observé à l'origine lors de l'introduction de transgènes partageant une homologie avec la séquence transcrite des gènes endogènes, mais il peut également se produire lors de l'introduction de transgènes ne présentant aucune homologie avec les séquences endogènes [38]. Ceci indique que pendant la co-suppression, les transgènes agissent comme des inducteurs de PTGS alors que les gènes endogènes se comportent uniquement comme des cibles. Encore une fois, l'intensité du PTGS des transgènes exogènes est corrélée avec le niveau d'ARN du transgène, ce qui soutient l'hypothèse d'un seuil. Ceci pose toutefois plusieurs questions : comment la cellule compte-t-elle les ARN, y compris ceux qui n'existent pas naturellement dans la plante? et quel type d'ARN sont comptés?

Une caractéristique frappante du PTGS des transgènes est son aspect dynamique et systémique, et son

effacement pendant la maturation des graines. En effet, les graines sèches ne présentent pas de PTGS, lequel commence après l'imbibition [39]. Une fois déclenché, il perdure jusqu'à la fin de la vie de la plante, et les plantes transmettent à leur descendance leur capacité à déclencher le PTGS, et non leur état de PTGS. Par conséquent, le PTGS se déclenche à chaque génération. Certaines lignées déclenchent le PTGS très rapidement après la germination, suggérant qu'il se produit de manière synchrone dans chaque cellule. Cependant, l'étude de lignées qui déclenchent le PTGS tardivement a révélé que le PTGS est initié localement, puis se propage à travers le système vasculaire pour finalement envahir de manière systémique la plante entière [40] (Figure 3A). Des expériences de greffage impliquant des greffons transgéniques actifs et des porte-greffes transgéniques éteints ont révélé qu'un signal séquence-spécifique est transmis des tissus éteints aux tissus actifs (Figure 3B et C), expliquant la propagation naturelle du PTGS dans toute la plante [41]. Cette propagation a été observée dans les cas de co-suppression (lignées p35S-NIA et p35S-NII, voir Figure 3) mais aussi dans les lignées portant un transgène p35S-GUS qui ne partage aucune homologie avec le génome de la plante. L'aspect systémique du PTGS a été rapidement confirmé en introduisant localement des copies supplémentaires du transgène p35S-GFP dans une plante transgénique p35S-GFP active et en observant le déclenchement du PTGS du transgène GFP dans les tissus systémiques [42]. L'analyse d'événements de co-suppression impliquant un transgène homologue aux gènes endogènes NIA ou NII a révélé que les greffons transgéniques actifs peuvent subir le PTGS lorsqu'ils sont greffés sur des porte-greffes transgéniques éteints, mais pas les greffons de type sauvage portant uniquement les gènes endogènes, ce qui indique que les gènes endogènes et les transgènes ont une compétence différente pour le PTGS [43] (Figure 3D). L'incapacité des gènes endogènes à subir un PTGS systémique en l'absence de transgènes homologues est abordée plus loin.

Pour déchiffrer le mécanisme moléculaire qui sous-tend le PTGS des transgènes, des cribles génétiques ont été mis en place afin d'identifier des mutants qui suppriment ou augmentent l'efficacité du PTGS. L'analyse de ces mutants a permis de dresser un modèle où le PTGS des transgènes est divisé en



**FIGURE 3.** Le PTGS transgénique est systémique et séquence-spécifique. (A) Établissement progressif du PTGS pendant le développement somatique. La co-suppression des gènes endogènes du NIA du tabac, induite par l'introduction de transgènes homologues, est visualisée par une chlorose foliaire. A chaque génération, le phénomène est initié localement sur une feuille. Puis, il envahit progressivement toute la plante en se propageant par les tissus vasculaires. Le même schéma est observé pour chaque lignée transgénique NIA. La co-suppression des gènes endogènes NII du tabac, induite par l'introduction de transgènes homologues, se produit de la même manière. (B) Le PTGS est transmis des porte-greffes inactivés par PTGS aux scions actifs (NS) portant le même transgène (à gauche : plantes p35S-NIA, à droite : plantes p35S-NII). (C) Le PTGS systémique est séquence-spécifique. Le PTGS n'est pas transmis des porte-greffes p35S-NII inactivés par PTGS aux scions p35S-NIA actifs (gauche) ou des porte-greffes p35S-NIA inactivés par PTGS aux scions p35S-NII actifs (droite). (D) Le PTGS systémique ne se produit qu'en présence d'un transgène. Les scions de type sauvage portant les gènes endogènes NIA et NII mais pas de transgènes homologues ne déclenchent pas de PTGS lorsqu'ils sont greffés sur des porte-greffes p35S-NIA (à gauche) ou p35S-NII (à droite) inactivés par PTGS.

quatre étapes : l'initiation primaire, l'amplification, la propagation et la ré-initiation qui conduit au PTGS systémique (Figure 4).

#### 4.1. *Initiation primaire*

Cette étape est encore assez mystérieuse. Néanmoins, il existe un consensus sur le fait que les ARN aberrants (ARNab) produits par les transgènes sont à l'origine du processus. Tout d'abord, parce que les mutations dans la voie du contrôle de la qualité de l'ARN (en anglais *RNA quality control* ou RQC), qui dégrade les ARNab, renforcent le PTGS [44–50], et ensuite parce qu'il a été démontré que les ARN non coiffés ou non polyadénylés sont des substrats préférentiels pour les RDR cellulaires, en particulier RDR6 [51–53]. Par conséquent, on suppose que RDR6 transforme les ARNab des transgènes en ARNdb pour initier le PTGS. Il est probable que ce processus se produise dans un nombre limité de cellules, soit en raison d'une sur-transcription accidentelle du transgène, ce qui entraîne des quantités d'ARNab transgéniques qui dépassent la capacité de dégradation du RQC, soit en raison d'une défaillance accidentelle du fonctionnement du RQC.

#### 4.2. *Amplification*

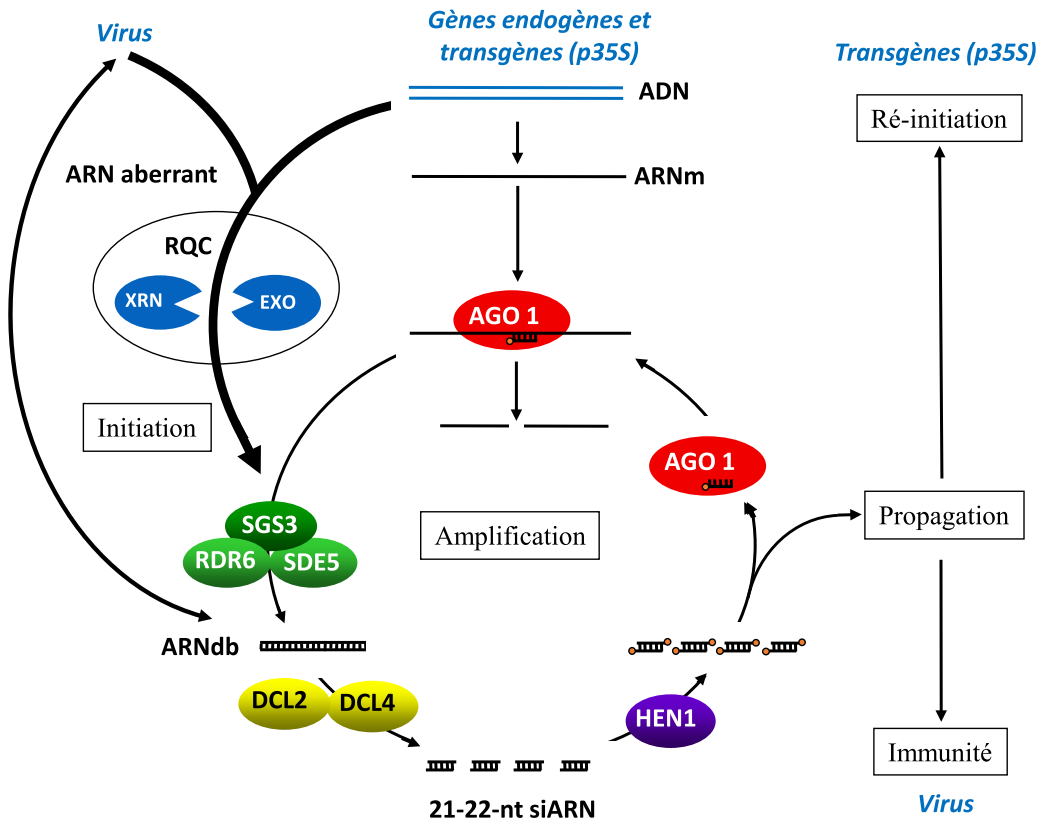
Les enzymes cytoplasmiques de type RNase III DCL4 et DCL2 coupent l'ARNdb en duplex de siARN de 21 et 22 nt, respectivement [54, 55]. Ces duplex sont méthylés par HEN1 à leurs extrémités 3' débordantes pour les protéger de l'uridylation et de la dégradation [56]. Ensuite, les duplex de siARN sont chargés dans des protéines de la famille ARGONAUTE (AGO), principalement AGO1 [57, 58]. Le brin passager du duplex de siARN est clivé et éliminé, ce qui permet l'hybridation du brin de siARN guide avec les molécules d'ARNm complémentaires du transgène et leur clivage grâce à l'activité RNaseH des protéines AGO. Après le clivage par les complexes AGO/siARN, les fragments d'ARNm transgénique sont dégradés par les exonucléases. Cependant, l'ARNm du transgène lié à un complexe AGO1/22-nt siRNA peut être transformé en ARNdb par RDR6, ce qui conduit à la production de siARN secondaires [59, 60] et à leur chargement ultérieur dans les protéines AGO, ce qui renforce l'élimination de l'ARNm du transgène.

#### 4.3. *Propagation*

Le PTGS devient finalement systémique en raison d'un signal qui peut se déplacer de cellule à cellule et à longue distance à travers les tissus vasculaires. Bien que la nature systémique du PTGS ait été découverte il y a 25 ans [41, 42], on sait encore peu de choses sur l'étape de propagation. En fait, on se demande encore si le signal est un ARNdb ou un siARN. Le mouvement des siARN a été clairement démontré [61, 62], mais la possibilité que le signal soit constitué d'ARNdb n'a pas été formellement écartée. De plus, on ne sait pas si l'ARN signal se déplace seul ou associé à des protéines cargo, et on ignore ce qui régule ce mouvement. Jusqu'à présent, les mutants altérés dans l'étape de propagation de l'ARN signal n'ont pas été identifiés, probablement parce que les éléments impliqués dans le mouvement de l'ARN signal sont essentiels pour la plante.

#### 4.4. *Ré-initiation*

Le PTGS devient systémique non seulement en raison de l'existence d'un ARN signal mobile mais aussi parce que les transgènes sont capables de ré-initier le PTGS dans les cellules où il n'a pas été initié par la saturation ou la déficience en RQC. Cette étape est également mal comprise, principalement parce qu'elle nécessite la réalisation d'expériences de greffe réciproque (tissus receveurs de type sauvage greffés sur des tissus émetteurs mutants vs tissus receveurs mutants greffés sur des tissus émetteurs de type sauvage) pour déterminer à quelle étape une mutation donnée entrave le PTGS systémique. Comme prévu, les mutations d'AGO1, de DCL2/DCL4, de RDR6 et de SGS3 dans les tissus émetteurs ou récepteurs inhibent le PTGS systémique, car il s'agit de protéines de base nécessaires dans chaque cellule pour accomplir le PTGS [63–66]. Jusqu'à présent, aucune mutation affectant spécifiquement la capacité des tissus émetteurs à transmettre le PTGS aux tissus récepteurs n'a été identifiée. En revanche, plusieurs mutations affectant spécifiquement la capacité des tissus récepteurs à déclencher le PTGS lorsqu'ils sont greffés sur des tissus émetteurs ont été identifiées. Cependant, le rôle des gènes correspondants reste énigmatique. En effet, on trouve des gènes codant pour une partie de la machinerie de TGS : PolIV, RDR2, DCL3, AGO4 [63, 65], et deux gènes, JM14 et NAC52,



**FIGURE 4.** Modèle de PTGS. Les ARN aberrants sont considérés comme étant à la base de la PTGS. Ils sont naturellement produits par les gènes endogènes, les transgènes et les virus, et devraient être éliminés par les voies de contrôle de la qualité des ARN (RQC). Cependant, lorsque, dans une cellule, la quantité d'ARN aberrants dépasse la capacité du RQC ou lorsque le RQC devient accidentellement dysfonctionnel, les ARN aberrants sont convertis en ARNdb par RDR6 pour initier le PTGS. Dans le cas des virus à ARN, les intermédiaires de la réplication virale sont constitués d'ARNdb qui peuvent aussi potentiellement initier la PTGS. Les enzymes cytoplasmiques RNase de type III DCL4 et DCL2 coupent l'ARNdb en duplex de siARN de 21 et 22 nt, qui sont méthylés par HEN1. Les ARNm complémentaires sont soit clivés par les complexes AGO1/siARN, soit copiés par RDR6 assisté de la protéine de liaison à l'ARN SGS3 et de la protéine d'exportation de l'ARN SDE5 pour produire davantage d'ARNdb, ce qui entraîne la production de siARN secondaires, maximisant ainsi l'élimination des ARNm. Le PTGS finit par se propager grâce à un signal ARN mobile qui peut se déplacer de cellule à cellule et à longue distance dans les tissus vasculaires. Il est probable que le signal ARN soit constitué de siARN, dont la mobilité a été clairement démontrée. Dans le cas des virus, si les siARN se propagent plus vite que le virus lui-même, ils confèrent une immunité aux tissus non infectés. Dans le cas des transgènes pilotés par le promoteur 35S, le PTGS se propage à toute la plante car les transgènes sont capables de réinitialiser le PTGS dans les cellules où il n'a pas été initié par la saturation ou la déficience en RQC. Cette capacité nécessite la déméthylase de l'histone H3K4me3 JM14 et le facteur de transcription à domaine NAC NAC52, qui protègent les séquences transgéniques de la méthylation de l'ADN *de novo*. Les séquences endogènes ne sont pas protégées de la méthylation de l'ADN *de novo* par JM14 et NAC52, et ne sont pas capables de subir une PTGS systémique, ce qui explique probablement comment les plantes survivent à une déficience locale et occasionnelle en RQC.

codant pour une histone H3K4me3 déméthylase et un facteur de transcription à domaine NAC. JMJ14 et NAC52 ne sont pas nécessaires à l'exécution du PTGS ni à la production de l'ARN signal, mais sont nécessaires pour que le PTGS du transgène devienne systémique [64, 67]. De façon remarquable, la méthylation de l'ADN augmente au niveau des séquences transgéniques en l'absence de JMJ14 ou de NAC52, mais pas au niveau des séquences endogènes [64]. Il est probable que ce n'est pas l'absence de JMJ14 ou de NAC52 mais l'hyperméthylation du transgène qui empêche la ré-initiation du PTGS. En effet, la ré-initiation du PTGS se produit normalement dans un fond mutant *jmj14* lorsque l'hyperméthylation est empêchée en mutant simultanément les ADN méthylases DRM2 et CMT3 [64]. La manière dont JMJ14 atténue spécifiquement la méthylation de l'ADN du transgène et favorise le PTGS systémique du transgène n'est toujours pas comprise.

## 5. Phénomènes naturels de PTGS

Le mécanisme à l'œuvre lorsque les plantes se remettent d'une infection virale est très similaire à celui impliqué dans le PTGS des transgènes (Figure 4). Il est probable que l'énorme quantité d'ARN viraux produite dans les premières cellules infectées comprenne suffisamment d'ARNab pour saturer le RQC, conduisant à la production d'ARNdb primaires qui sont traités dans le cytoplasme par DCL4 et DCL2, aboutissant à des duplex de siARN primaires de 21 et 22 nt qui sont chargés sur des protéines de la famille ARGONAUTE (AGO), principalement AGO1, mais aussi AGO2, AGO5, AGO7 et AGO10 [68]. Alternativement, ou concomitamment, les ARNdb viraux formés pendant la réplication du virus pourraient être directement utilisés par les DCL pour produire les siARN primaires. Une fois que les siARN sont chargés dans les AGO, ils guident le clivage des grands ARN viraux. De manière similaire à ce qui se passe lors du PTGS du transgène, le complexe AGO1/22-nt siARN favorise spécifiquement la production d'ARNdb supplémentaires par RDR6, conduisant à la production de siARN secondaires et à leur chargement ultérieur sur les protéines AGO, ce qui renforce l'élimination des ARN viraux de la cellule végétale.

Alors que les transgènes dirigés par le promoteur viral 35S sont souvent éteints par PTGS, les exemples

de transgènes dirigés par des promoteurs endogènes et subissant le PTGS manquent. Néanmoins, il a été démontré qu'une poignée de gènes codant pour des protéines (en anglais : *protein-coding gene* ou PCG), dans certaines variétés, produisent des siARN à des niveaux qui permettent la dégradation des ARNm par PTGS. Ces variétés présentent des réarrangements génomiques impliquant des événements de duplication, qui permettent de produire des ARNdb, alors que les variétés normales présentent une seule copie de gène qui ne produit pas de siARN [7, 69–72]. Il est probable que de tels réarrangements sont tolérés parce qu'ils affectent des gènes non indispensables à la survie de la plante. Cela implique que (i) au cours de l'évolution, les gènes codant pour les protéines endogènes ont été façonnés pour éviter de produire des quantités excessives de siARN, (ii) les cellules subissant des événements de duplication qui provoquent la production de siARN à partir de gènes essentiels ne survivent pas. Il est important de noter que les événements de PTGS impliquant des transgènes sont toujours systémiques, alors que les quelques exemples connus de PCG subissant le PTGS ne montrent qu'une extinction localisée, confirmant ce qui a été observé en greffant des plantes de type sauvage sur des plantes transgéniques subissant le PTGS [43].

Les résultats décrits ci-dessus suggèrent que seuls les réarrangements génomiques conduisant à la production d'ARNdb pouvaient favoriser la production de siRNA à partir de PCG. Cependant, la plupart des mutants déficients en RQC sont connus depuis longtemps pour mourir à un stade très précoce de leur développement, ce qui suggère que les ARNab produits par les PCG pourraient entrer dans la voie de la PTGS lorsqu'ils ne sont pas éliminés par le RQC, permettant ainsi de produire des siARN qui pourraient cibler des ARNm homologues fonctionnels pour les détruire, ce qui finit par tuer la plante. À l'appui de cette hypothèse, la létalité des mutants déficients en RQC est levée quand on mute simultanément la voie PTGS [50, 73]. L'inventaire des siARN dans des mutants d'*Arabidopsis* altérés dans les activités RQC a révélé que des milliers de PCG produisent des siARN (appelés *rqc-siARN*) lorsque les plantes sont déficientes pour RQC. Étant donné que les allèles nuls RQC sont embryolétaux ou meurent à un stade très précoce du développement, il est impossible d'obtenir une vue complète des PCG capables

de produire des siARN en utilisant des mutants totalement déficients pour le RQC. L'analyse de mutants qui altèrent partiellement le RQC a permis d'étendre la liste des PCG produisant des siARN [50, 73–75]. De plus, une autre catégorie de siARN dérivés des PCG a été identifiée lorsque les plantes sont infectées par des virus [76]. Ces siARN activés par des virus ont été appelés va-siARN. Il existe un overlap important entre les rqc-siARN et les va-siARN, ce qui suggère que les virus pourraient provoquer la production de siARN à partir des PCG en exprimant des protéines qui inhibent le RQC ou en produisant des ARNab jusqu'à un niveau qui sature la voie RQC et déclenche l'entrée des ARNab dérivés des PCG dans la voie PTGS. À ce jour, on a constaté que ~5000 PCG parmi les ~27 000 PCG d'*Arabidopsis* étaient capables de produire des siARN. On ignore encore si seuls ces ~5000 PCG sont capables de produire des siRNAs ou si leur nombre est sous-estimé en raison de l'analyse restreinte des tissus survivant dans les plantes mutantes déficientes en RQC ou infectées par des virus.

Il est possible qu'au cours de l'évolution, le RQC ait été ajusté à la quantité d'ARNab produits par les PCG endogènes de sorte que les PCG ne produisent pas de siARN dans les cellules déficientes en RQC. Il est également possible que la déficience locale ou transitoire du RQC n'ait pas de conséquences dramatiques sur la plante entière car, contrairement aux transgènes, les PCG endogènes sont incapables de déclencher un PTGS systémique. Cela pourrait être lié à la régulation épigénétique distincte des transgènes et des PCG endogènes. En effet, alors que la méthylation de l'ADN des PCG endogènes est limitée par l'histone H3K9me2 déméthylase IBM1, la méthylation de l'ADN des transgènes exprimés sous le contrôle du promoteur viral 35S est limitée par l'histone H3K4me3 déméthylase JMJ14 mais pas par IBM1. JMJ14 n'est pas nécessaire pour l'exécution du PTGS ni pour la production de l'ARN signal, mais il est nécessaire pour que le PTGS des transgènes devienne systémique [64]. La façon dont JMJ14 atténue spécifiquement la méthylation de l'ADN du transgène et favorise la PTGS systémique du transgène n'est toujours pas comprise. Néanmoins, les régulations épigénétiques distinctes des transgènes et des PCG endogènes impliquent qu'une production locale de siARN endogènes n'a aucune conséquence sur l'expression des PCG endogènes dans le reste de

la plante, alors qu'une production locale de siARN transgéniques active le PTGS systémique du transgène. Dans le modèle actuel du PTGS, une fois que les siARN sont produits dans certaines cellules, ils se déplacent vers d'autres cellules mais seuls les transgènes, et non les PCG endogènes, peuvent ré-initier le PTGS dans les cellules recevant les siARN en raison de la régulation spécifique des transgènes par JMJ14. Par conséquent, les transgènes exprimés sous le contrôle du promoteur 35S sont susceptibles de produire des siARN de 21 et 22 nt et de déclencher le PTGS systémique, alors que les PCG endogènes produisant des siRNA de 21 et 22 nt et présentant des signes de PTGS sont rares et ne montrent qu'un PTGS localisé [7, 69–72].

## 6. Phénomènes naturels impliquant à la fois TGS et PTGS

Dans les sections précédentes, le TGS et le PTGS ont été définis comme des phénomènes indépendants, agissant sur des cibles différentes. Bien que ce soit généralement le cas, les éléments transposables se distinguent, impliquant TGS et PTGS aux différentes étapes de leur contrôle épigénétique. Le maintien des transposons dans un état inactif implique le mécanisme RdDM/TGS décrit dans la Figure 2. Cependant, malgré ce mécanisme de maintien [18], les transposons peuvent parfois être réactivés, par exemple dans le pollen où DDM1 n'est pas exprimé [77], ou dans des mutants où DDM1 ou MET1 sont artificiellement altérés [78, 79]. Dans ces cas, les transposons sont d'abord soumis au PTGS, avant d'être éteints par TGS. On suppose que les transcrits des transposons subissent soit un blocage des ribosomes, ce qui provoque un clivage de l'ARN [80], soit sont clivés par des miRNA spécifiques, ce qui permet aux produits de clivage d'être transformés en dsRNA par RDR6 [81]. Ces ARNdb dérivés de transposons sont transformés par DCL4 et DCL2 en siARN de 21 et 22 nt, qui sont chargés dans AGO1 et AGO2. Cependant, le PTGS est insuffisant pour supprimer l'expression de transposons. C'est l'action de DCL3 sur les ARNdb dérivés des transposons qui provoque la production de siARN de 24 nt qui sont chargés dans les AGO de type AGO4, permettant de guider la méthylation de l'ADN vers les copies réactivées de transposons, ce qui entraîne la ré-initiation du TGS [82], qui peut

ensuite être maintenue par la voie canonique RdDM (Figure 2).

## 7. Remarques finales

Les génomes sont en constante évolution. À chaque génération, les *crossing-overs*, obligatoires pour que la méiose se déroule correctement [83], juxtaposent des séquences qui n'étaient pas adjacentes dans les génomes parentaux. Les croisements entre différentes plantes assemblent dans leur descendance des génomes qui peuvent fortement différer par leurs compositions en longs et petits ARN non codants ainsi que par les marques épigénétiques qui leur sont associées [84, 85]. Lorsqu'ils sont soumis à des stress génotoxiques, les génomes peuvent être endommagés, et la réparation qui s'ensuit peut être parfaite ou imparfaite. Les génomes peuvent également subir une duplication partielle ou complète [86]. Les transposons peuvent se déplacer ou se dupliquer, et les nouvelles copies insérées peuvent perturber un gène ou s'insérer à proximité et modifier son expression [87]. Enfin, certains pathogènes peuvent insérer une partie de leur matériel génétique dans le génome de la plante, et avoir des conséquences similaires aux insertions de transposons. Pour toutes ces raisons, il est essentiel pour la cellule de distinguer le soi du soi modifié/endommagé et du non-soi. La séquence d'ADN est clairement insuffisante pour fournir cette information. Il est probable que les marques épigénétiques qui décorent la chromatine permettent ce processus d'identification. Mieux, ce pourrait être simplement l'absence de marques épigénétiques natives qui permet d'identifier le non-soi ou le soi modifié/endommagé. En effet, les nouvelles séquences d'ADN (ADN transgénique, ADN pathogène, nouvelle copie d'élément transposable, ADN endommagé/réparé, fragment génomique dupliqué) sont généralement nues. Pour former la chromatine, l'ADN s'assemble avec des histones, mais comment sont choisies les marques épigénétiques qui décorent les histones qui s'associent au nouvel ADN? Par défaut, des histones portant des décorations particulières pourraient s'associer à de l'ADN nu, marquant ces nouvelles séquences comme potentiellement dangereuses, nécessitant une extinction immédiate, ou au moins une surveillance pendant une longue période avant qu'elles ne soient considérées comme sûres et faisant donc partie intégrante du génome.

## Conflit d'intérêt

L'auteur n'a aucun conflit d'intérêt à déclarer.

## Remerciements

Tout d'abord, je tiens à remercier Michel Caboche pour m'avoir accueilli dans son laboratoire en 1985 et m'avoir transmis sa passion pour la génétique végétale. Son mentorat a été essentiel pour ma carrière. Je remercie également Jean-Pierre Bourgin, Francine Casse, Yves Chupeau, Georges Pelletier et Pierre Rouzé pour leur soutien et leurs encouragements à une époque où l'extinction de gènes était encore un mystère. Merci à Jean-Luc Rossignol et Godelaine Faugeron d'avoir partagé leurs idées sur l'épigénétique alors que la plupart des gens ne savaient pas ce que c'était. Un grand merci aux membres de mon groupe pour avoir supporté mon intérêt pour les ségrégations aberrantes, et en particulier Taline Elmayan pour 30 ans de collaboration fidèle et enthousiaste. Enfin, je tiens à remercier les membres de la communauté internationale de l'extinction des gènes, et en particulier Marjori et Antonius Matzke pour les discussions fructueuses autour des premières plantes transgéniques.

## References

- [1] P. Cubas, C. Vincent, E. Coen, "An epigenetic mutation responsible for natural variation in floral symmetry", *Nature* **401** (1999), p. 157-161.
- [2] K. C. Sink, "Taxonomy", in *Petunia* (K. C. Sink, ed.), Springer, New York, 1984, p. 3-9.
- [3] J. R. Stehmann, A. P. Lorenz-Lemke, L. B. Freitas, J. Semir, "The genus *Petunia*", in *Petunia* (T. Gerats, J. Strommer, eds.), Springer, New York, 2009, p. 1-28.
- [4] L. H. Bailey, "Evolution of the *Petunia*", in *The Survival of the Unlike*, MacMillan, New York, 1896, p. 465-472.
- [5] C. Napoli, C. Lemieux, R. Jorgensen, "Introduction of a chimeric chalcone synthase gene into *petunia* results in reversible co-suppression of homologous gene *in trans*", *Plant Cell* **2** (1990), p. 279-289.
- [6] A. R. van der Krol, L. A. Mur, M. Beld, J. N. Mol, A. R. Stuitje, "Flavonoid genes in *petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression", *Plant Cell* **2** (1990), p. 291-299.
- [7] Y. Morita *et al.*, "Tandemly arranged chalcone synthase A genes contribute to the spatially regulated expression of siRNA and the natural bicolor floral phenotype in *Petunia hybrida*", *Plant J.* **70** (2012), p. 739-749.
- [8] S. A. Wingard, "Hosts and symptoms of ring spot, a virus disease of plants", *J. Agric. Res.* **37** (1928), p. 127-153.



- [9] D. Baulcombe, "RNA silencing in plants", *Nature* **431** (2004), p. 356-363.
- [10] K. A. Barton, A. N. Binns, A. J. Matzke, M. D. Chilton, "Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny", *Cell* **32** (1983), p. 1033-1043.
- [11] M. A. Matzke, M. Primig, J. Trnovsky, A. J. Matzke, "Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants", *Embo J.* **8** (1989), p. 643-649.
- [12] Y. D. Park *et al.*, "Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity", *Plant J.* **9** (1996), p. 183-194.
- [13] G. Faugeron, C. Goyon, A. Gregoire, "Stable allele replacement and unstable non-homologous integration events during transformation of *Ascomobolus immersus*", *Gene* **76** (1989), p. 109-119.
- [14] F. Linn, I. Heidmann, H. Saedler, P. Meyer, "Epigenetic changes in the number of the maize A1 gene in *Petunia* hybrids: role of numbers of integrated gene copies and state of methylation", *Mol. Gen. Genet.* **222** (1990), p. 329-336.
- [15] H. Vaucheret, "Identification of a general silencer for 19S and 35S promoter in transgenic tobacco: 90 pb of homology in the promoter sequences are sufficient for trans-inactivation", *C. R. Acad. Sci. Paris, Life Sci.* **316** (1993), p. 1471-1483.
- [16] M. F. Mette, W. Aufsatz, J. van der Winden, M. A. Matzke, A. J. Matzke, "Transcriptional silencing and promoter methylation triggered by double-stranded RNA", *Embo J.* **19** (2000), p. 5194-5201.
- [17] P. Mourrain, R. van Blokland, J. M. Kooter, H. Vaucheret, "A single transgene locus triggers both transcriptional and post-transcriptional silencing through double-stranded RNA production", *Planta* **225** (2007), p. 365-379.
- [18] M. A. Matzke, T. Kanno, A. J. Matzke, "RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants", *Annu. Rev. Plant Biol.* **66** (2015), p. 243-267.
- [19] C. Eun *et al.*, "Use of forward genetic screens to identify genes required for RNA-directed DNA methylation in *Arabidopsis thaliana*", *Cold Spring Harb. Symp. Quant. Biol.* **77** (2012), p. 195-204.
- [20] Z. Gong *et al.*, "ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase", *Cell* **111** (2002), p. 803-814.
- [21] X. Cao, S. E. Jacobsen, "Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing", *Curr. Biol.* **12** (2002), p. 1138-1144.
- [22] S. W. Chan *et al.*, "RNA silencing genes control de novo DNA methylation", *Science* **303** (2004), p. 1336.
- [23] M. A. Arteaga-Vazquez, V. L. Chandler, "Paramutation in maize: RNA mediated trans-generational gene silencing", *Curr. Opin. Genet. Dev.* **20** (2010), p. 156-163.
- [24] R. A. Brink, "Paramutation at the R locus in maize", *Cold Spring Harb. Symp. Quant. Biol.* **23** (1958), p. 379-391.
- [25] J. B. Hollick, "Paramutation and related phenomena in diverse species", *Nat. Rev. Genet.* **18** (2017), p. 5-23.
- [26] M. Stam *et al.*, "The regulatory regions required for B' paramutation and expression are located far upstream of the maize b1 transcribed sequences", *Genetics* **162** (2002), p. 917-930.
- [27] J. Bender, G. R. Fink, "Epigenetic control of an endogenous gene family is revealed by a novel blue fluorescent mutant of *Arabidopsis*", *Cell* **83** (1995), p. 725-734.
- [28] S. Melquist, J. Bender, "Transcription from an upstream promoter controls methylation signaling from an inverted repeat of endogenous genes in *Arabidopsis*", *Genes Dev.* **17** (2003), p. 2036-2047.
- [29] S. Durand, N. Bouche, E. Perez Strand, O. Loudet, C. Camilleri, "Rapid establishment of genetic incompatibility through natural epigenetic variation", *Curr. Biol.* **22** (2012), p. 326-331.
- [30] A. Martin, C. Troadec, A. Boualem, M. Rajab, R. Fernandez, H. Morin, M. Pitrat, C. Dogimont, A. Bendahmane, "A transposon-induced epigenetic change leads to sex determination in melon", *Nature* **461** (2009), p. 1135-1138.
- [31] F. de Carvalho, G. Gheysen, S. Kushnir, M. Van Montagu, D. Inzé, C. Castresana, "Suppression of beta-1,3-glucanase transgene expression in homozygous plants", *EMBO J.* **11** (1992), p. 2595-2602.
- [32] F. Dorlhac de Borne, M. Vincentz, Y. Chupeau, H. Vaucheret, "Co-suppression of nitrate reductase host genes and transgenes in transgenic tobacco plants", *Mol. Gen. Genet.* **243** (1994), p. 613-621.
- [33] C. M. Hart, B. Fischer, J. M. Neuhaus, F. Meins Jr., "Regulated inactivation of homologous gene expression in transgenic *Nicotiana sylvestris* plants containing a defense-related tobacco chitinase gene", *Mol. Gen. Genet.* **235** (1992), p. 179-188.
- [34] C. J. Smith, C. F. Watson, C. R. Bird, J. Ray, W. Schuch, D. Grierson, "Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants", *Mol. Gen. Genet.* **224** (1990), p. 477-481.
- [35] J. C. Palauqui, H. Vaucheret, "Field trial analysis of nitrate reductase co-suppression: a comparative study of 38 combinations of transgene loci", *Plant Mol. Biol.* **29** (1995), p. 149-159.
- [36] H. Vaucheret, L. Nussbaum, J. C. Palauqui, I. Quillere, T. Elmayer, "A transcriptionally active state is required for post-transcriptional silencing (cosuppression) of nitrate reductase host genes and transgenes", *Plant Cell* **9** (1997), p. 1495-1504.
- [37] A. J. Hamilton, D. C. Baulcombe, "A novel species of small antisense RNA in posttranscriptional gene silencing", *Science* **286** (1999), p. 950-952.
- [38] T. Elmayer, H. Vaucheret, "Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally", *Plant J.* **9** (1996), p. 787-797.
- [39] T. Elmayer, S. Balzergue, F. Béon, V. Bourdon, J. Daubremet, Y. Guénet, P. Mourrain, J. C. Palauqui, S. Vernhettes, T. Vialle, K. Wostrikoff, H. Vaucheret, "Arabidopsis mutants impaired in cosuppression", *Plant Cell* **10** (1998), p. 1747-1758.
- [40] J. C. Palauqui, T. Elmayer, F. D. De Borne, P. Crete, C. Charles, H. Vaucheret, "Frequencies, timing, and spatial patterns of co-suppression of nitrate reductase and nitrite reductase in transgenic tobacco plants", *Plant Physiol.* **112** (1996), p. 1447-1456.
- [41] J. C. Palauqui, T. Elmayer, J. M. Pollien, H. Vaucheret, "Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from

- silenced stocks to non-silenced scions”, *Embo J.* **16** (1997), p. 4738-4745.
- [42] O. Voinnet, D. C. Baulcombe, “Systemic signalling in gene silencing”, *Nature* **389** (1997), p. 553.
- [43] J. C. Palauqui, H. Vaucheret, “Transgenes are dispensable for the RNA degradation step of cosuppression”, *Proc. Natl Acad. Sci. USA* **95** (1998), p. 9675-9680.
- [44] S. Gazzani, T. Lawrenson, C. Woodward, D. Headon, R. Sablowski, “A link between mRNA turnover and RNA interference in Arabidopsis”, *Science* **306** (2004), p. 1046-1048.
- [45] I. Gy, V. Gascioli, D. Laressergues, J.-B. Morel, J. Gombert, F. Proux, C. Proux, H. Vaucheret, A. C. Mallory, “Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors”, *Plant Cell* **19** (2007), p. 3451-3461.
- [46] K. Hematy, Y. Bellec, R. Podicheti, N. Bouteiller, P. Anne, C. Morineau, R. P. Haslam, F. Beaudoin, J. A. Napier, K. Mockaitis, D. Gagliardi, H. Vaucheret, H. Lange, J.-D. Faure, “The zinc-finger protein SOP1 is required for a subset of the nuclear exosome functions in arabidopsis”, *PLoS Genet.* **12** (2016), article no. e1005817.
- [47] H. Lange *et al.*, “The RNA helicases AtMTR4 and HEN2 target specific subsets of nuclear transcripts for degradation by the nuclear exosome in Arabidopsis thaliana”, *PLoS Genet.* **10** (2014), article no. e1004564.
- [48] A. B. Moreno *et al.*, “Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants”, *Nucleic Acids Res.* **41** (2013), p. 4699-4708.
- [49] A. Yu *et al.*, “Second-site mutagenesis of a hypomorphic argonaute1 allele identifies SUPERKILLER3 as an endogenous suppressor of transgene posttranscriptional gene silencing”, *Plant Physiol.* **169** (2015), p. 1266-1274.
- [50] X. Zhang *et al.*, “Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis”, *Science* **348** (2015), p. 120-123.
- [51] K. Baeg, H. O. Iwakawa, Y. Tomari, “The poly(A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing”, *Nat. Plants* **3** (2017), article no. 17036.
- [52] T. Dalmay, A. Hamilton, S. Rudd, S. Angell, D. C. Baulcombe, “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** (2000), p. 543-553.
- [53] P. Mourrain *et al.*, “Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance”, *Cell* **101** (2000), p. 533-542.
- [54] V. Gascioli, A. C. Mallory, D. P. Bartel, H. Vaucheret, “Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs”, *Curr. Biol.* **15** (2005), p. 1494-1500.
- [55] J. S. Parent, N. Bouteiller, T. Elmayan, H. Vaucheret, “Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing”, *Plant J.* **81** (2015), p. 223-232.
- [56] Z. Yang, Y. W. Ebright, B. Yu, X. Chen, “HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide”, *Nucleic Acids Res.* **34** (2006), p. 667-675.
- [57] M. Fagard, S. Boutet, J. B. Morel, C. Bellini, H. Vaucheret, “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. USA* **97** (2000), p. 11650-11654.
- [58] J. B. Morel *et al.*, “Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance”, *Plant Cell* **14** (2002), p. 629-639.
- [59] H. O. Iwakawa *et al.*, “Ribosome stalling caused by the Argonaute-microRNA-SGS3 complex regulates the production of secondary siRNAs in plants”, *Cell Rep.* **35** (2021), article no. 109300.
- [60] Y. Sakurai *et al.*, “Cell-free reconstitution reveals the molecular mechanisms for the initiation of secondary siRNA biogenesis in plants”, *Proc. Natl Acad. Sci. USA* **118** (2021), no. 31, article no. e21028891.
- [61] C. W. Melnyk, A. Molnar, A. Bassett, D. C. Baulcombe, “Mobile 24 nt small RNAs direct transcriptional gene silencing in the root meristems of Arabidopsis thaliana”, *Curr. Biol.* **21** (2011), p. 1678-1683.
- [62] A. Molnar *et al.*, “Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells”, *Science* **328** (2010), p. 872-875.
- [63] C. A. Brosnan *et al.*, “Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis”, *Proc. Natl Acad. Sci. USA* **104** (2007), p. 14741-14746.
- [64] N. Butel *et al.*, “Contrasting epigenetic control of transgenes and endogenous genes promotes post-transcriptional transgene silencing in Arabidopsis”, *Nat. Commun.* **12** (2021), article no. 2787.
- [65] C. Taochy *et al.*, “A genetic screen for impaired systemic RNAi highlights the crucial role of DICER-LIKE 2”, *Plant Physiol.* **175** (2017), p. 1424-1437.
- [66] C. Taochy *et al.*, “Post-transcriptional gene silencing triggers dispensable DNA methylation in gene body in Arabidopsis”, *Nucleic Acids Res.* **47** (2019), p. 9104-9114.
- [67] I. Le Masson *et al.*, “Mutations in the Arabidopsis H3K4me2/3 demethylase MJM14 suppress posttranscriptional gene silencing by decreasing transgene transcription”, *Plant Cell* **24** (2012), p. 3603-3612.
- [68] A. Carbonell, J. C. Carrington, “Antiviral roles of plant ARGONAUTES”, *Curr. Opin. Plant Biol.* **27** (2015), p. 111-117.
- [69] S. J. Clough *et al.*, “Features of a 103-kb gene-rich region in soybean include an inverted perfect repeat cluster of CHS genes comprising the I locus”, *Genome* **47** (2004), p. 819-831.
- [70] C. B. Della Vedova *et al.*, “The dominant inhibitory chalcone synthase allele C2-Idf (inhibitor diffuse) from Zea mays (L.) acts via an endogenous RNA silencing mechanism”, *Genetics* **170** (2005), p. 1989-2002.
- [71] J. H. Tuteja, S. J. Clough, W. C. Chan, L. O. Vodkin, “Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in Glycine max”, *Plant Cell* **16** (2004), p. 819-835.
- [72] J. H. Tuteja, G. Zabala, K. Varala, M. Hudson, L. O. Vodkin, “Endogenous, tissue-specific short interfering RNAs silence the chalcone synthase gene family in glycine max seed coats”, *Plant Cell* **21** (2009), p. 3063-3077.
- [73] A. E. Martinez de Alba *et al.*, “In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs”, *Nucleic Acids Res.* **43** (2015), p. 2902-2913.
- [74] M. Krzysztoson, J. Kufel, “Analysis of mRNA-derived siRNAs in

- mutants of mRNA maturation and surveillance pathways in *Arabidopsis thaliana*", *Sci. Rep.* **12** (2022), article no. 1474.
- [75] H. Scheer *et al.*, "The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis", *Nat. Commun.* **12** (2021), article no. 1298.
- [76] M. Cao *et al.*, "Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*", *Proc. Natl Acad. Sci. USA* **111** (2014), p. 14613-14618.
- [77] R. K. Slotkin *et al.*, "Epigenetic reprogramming and small RNA silencing of transposable elements in pollen", *Cell* **136** (2009), p. 461-472.
- [78] F. Johannes *et al.*, "Assessing the impact of transgenerational epigenetic variation on complex traits", *PLoS Genet.* **5** (2009), article no. e1000530.
- [79] J. Reinders *et al.*, "Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes", *Genes Dev.* **23** (2009), p. 939-950.
- [80] S. Oberlin *et al.*, "Innate, translation-dependent silencing of an invasive transposon in *Arabidopsis*", *EMBO Rep.* **23** (2022), article no. e53400.
- [81] K. M. Creasey *et al.*, "miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*", *Nature* **508** (2014), p. 411-415.
- [82] A. Mari-Ordonez *et al.*, "Reconstructing de novo silencing of an active plant retrotransposon", *Nat. Genet.* **45** (2013), p. 1029-1039.
- [83] R. Mercier, C. Mezard, E. Jenczewski, N. Macaisne, M. Grelon, "The molecular biology of meiosis in plants", *Annu. Rev. Plant Biol.* **66** (2015), p. 297-327.
- [84] P. V. Shivaprasad, R. M. Dunn, B. A. Santos, A. Bassett, D. C. Baulcombe, "Extraordinary transgressive phenotypes of hybrid tomato are influenced by epigenetics and small silencing RNAs.", *EMBO J.* **31** (2012), p. 257-266.
- [85] M. W. Vaughn *et al.*, "Epigenetic natural variation in *Arabidopsis thaliana*", *PLoS Biol.* **5** (2007), article no. e174.
- [86] N. Panchy, M. Lehti-Shiu, S. H. Shiu, "Evolution of gene duplication in plants", *Plant Physiol.* **171** (2016), p. 2294-2316.
- [87] C. Mhiri, F. Borges, M. A. Grandbastien, "Specificities and dynamics of transposable elements in land plants", *Biology (Basel)* **11** (2022), no. 4, article no. 488.