

CURRENT REVIEW

A Four-Partner Plant–Virus Interaction: Enemies Can Also Come from Within

Marie-Line Iskra-Caruana,¹ Franc-Christophe Baurens,² Philippe Gayral,³ and Matthieu Chabannes¹

¹CIRAD, UMR BGPI, and ²CIRAD, UMR DAP, F-34398 Montpellier, France; ³Institut des Sciences de l'Evolution, CNRS UMR 5554, Université Montpellier 2, Place E. Bataillon, 34095 Montpellier, France

Submitted 10 May 2010. Accepted 21 June 2010.

Plant viruses are disseminated by either vertical (vegetative multiplication or sexual reproduction) or horizontal (vector-mediated) propagation. Plant pararetroviruses—members of the *Caulimoviridae* family—have developed an alternative strategy for vertical propagation via integration within the host plant genome, although integration is not required for viral replication. Integrated endogenous pararetrovirus (EPRV) sequences have undergone extensive viral genome rearrangements and contain more than one copy of the viral genome. Furthermore, EPRV can become infectious upon spontaneous escape of active virus following stresses such as wounding, tissue culture, or interspecific crosses. Such infectious EPRV are of great importance, not only in terms of their ability to precipitate epidemic outbreaks but also because of their effect on breeding of numerous plant genomes in temperate and tropical crops. This is especially true for banana, a crop susceptible to banana streak viruses, the causative agents of banana streak disease. Thus, the classical three-component banana–*Banana streak virus* (BSV)—mealybug pathosystem can be expanded to include endogenous BSV as an alternative source of active virions. The BSV-banana pathosystem is one of only three pathosystems known to date to harbor this remarkable feature, and the present review focuses exclusively on it to illustrate this four-partner interaction.

Until fairly recently, one of the basic tenets of virology was that plant viruses, unlike animal and bacterial viruses, did not integrate into the host genome. Observations over the past decade have broken this tenet, and an increasing number of integrated plant viral sequences have now been found in plant genomes (Harper et al. 2002; Hull et al. 2000; Hohn et al. 2008). Indeed, sequences derived from both groups of plant DNA viruses—the single-stranded DNA geminiviruses (Ashby et al. 1997; Bejarano et al. 1996; Kenton et al. 1995) and the double-stranded DNA pararetroviruses (Budiman et al. 2000; Gambley et al. 2008; Gayral et al. 2008; Harper et al. 1999, 2002; Jakowitsch et al. 1999; Lockhart et al. 2000; Mao et al. 2000; Ndowora et al. 1999; Pahalawatta et al. 2008; Staginnus and Richert-Poggeler 2006;)—have been found inserted into plant chromosomes. Integration of viral sequences has apparently occurred by illegitimate recombination (Jakowitsch et al. 1999) because the viruses involved do not encode an integrase function and do not require integration into the host genome as

part of their replication cycle. Like retroviruses, pararetroviruses use a reverse transcriptase for their replication. Pararetroviruses integrated into the host genome are referred to as endogenous pararetroviruses (EPRV) (Mette et al. 2002; Staginnus et al. 2009).

Plant pararetroviruses compose only one family, the *Caulimoviridae*, which contains six genera (*Caulimovirus*, *Soymovirus*, *Cavemovirus*, *Petuvirus*, *Badnavirus*, and *Tungrovirus*) (Fauquet et al. 2005; Hull 2001, 2002). To date, viruses from five of the six genera of this family (the exception being genus *Soymovirus*) have been reported to have integrants in plant genomes. Integrants identified to date include *Banana streak virus* (BSV) in *Musa* spp. (Gayral et al. 2008; Geering et al. 2001, 2005; Harper et al. 1999, 2005; Ndowora et al. 1999), *Petunia vein-clearing virus* (PVCV) in petunia (Harper et al. 2003; Noreen et al. 2007; Richert-Poggeler and Shepherd 1997; Richert-Poggeler et al. 2003), *Tobacco vein-clearing virus* (TVCV) in tobacco (Gregor et al. 2004; Jakowitsch et al. 1999; Lockhart et al. 2000), *Rice tungro bacilliform virus* (RTBV) in rice (Kunii et al. 2004; Liu et al. 2009), TVCV-like virus in tomato (Staginnus et al. 2007), pararetrovirus-like sequences in potato (Hansen et al. 2005), *Pineapple bacilliform virus* (PBV) in pineapple (Gambley et al. 2008), and *Dahlia mosaic caulimovirus* (DMV) in dahlia (Pahalawatta et al. 2008) genomes. The description of several other integrations in various plant species, such as passion fruit, citrus, and so on, is in progress (P.-Y. Teycheney et al. *unpublished data*).

However, with the exception of a few examples, such PVCV (Richert-Poggeler et al. 2003), TVCV (Lockhart et al. 2000), and distinct BSV species (*Banana streak OL virus* [BSOLV], *Banana streak Imové virus* [BSImV], and *Banana streak GF virus* [BSGFV]) (Gayral et al. 2008; Harper and Hull 1998; Harper et al. 2002; Ndowora et al. 1999; Safar et al. 2004), most reported plant EPRV are defective and cannot trigger virus infection.

The role of EPRV in plant–virus interactions is not yet fully understood. Current information suggests that EPRV are not always neutral components of plant genomes but can potentially contribute to either pathogenicity or virus resistance in the host. Indeed, Mette and associates (2002) and Noreen and associates (2007) provided the first evidence to support the view that EPRV could have beneficial effects for their hosts by providing virus resistance via homology-dependent transcriptional or posttranscriptional gene silencing.

Under different conditions or in other genome constitutions, these integrated viral sequences can be reactivated in response to abiotic or genomic stresses. Circular viral genome copies are probably formed by transcription of a more-than-genome-length

Corresponding author: M.-L. Iskra-Caruana;
E-mail: marie-line.caruana@cirad.fr

copy from tandemly arranged integrants (Richert-Poggeler et al. 2003) or by excision following recombination from fragmented integrants (Ndowora et al. 1999) (discussed below), which leads to the assembly of virus particles and symptoms of viral infection. Interspecific genetic crosses and in vitro propagation can induce EPRV activation, which has been shown to be economically detrimental in banana breeding (Dallot et al. 2001; Lheureux et al. 2003; Ndowora et al. 1999).

This review will focus on describing an inventory of knowledge of this novel and peculiar BSV–banana pathosystem.

Banana streak disease: a classic triangular pathogenic interaction.

Banana (*Musa* spp.) encompasses a wide range of dessert cultivars as well as cooking cultivars such as plantains. Banana fruit are the fourth most important food commodity in the world and represent a staple food for millions of people in developing countries as well as an important cash crop for many developing countries in South America, Southeast Asia, and Africa (Heslop-Harrison and Schwarzacher 2007). The majority of cultivated banana lines originated mainly from inter- and intraspecific crosses between the two wild species *Musa acuminata* (AA genome) and *M. balbisiana* (BB genome) (Simmonds and Shepherd 1955). A series of diploid, triploid, and tetraploid hybrid genomes (AA, AB, AAA, AAB, ABB, AABB, AAAB, and ABBB) resulting from such crosses are used in three types of food: as dessert banana fruit, in cooking (including plantains), and as beverages (beers). Wild *M. balbisiana* is endemic to drier areas of Southeast Asia (De Langhe et al. 2009; Perrier et al. 2009; Simmonds 1962), is extremely robust and fast growing, and is described as a source of resistance to various biotic and abiotic stresses (Bakry et al. 2009). In addition, the B genome constitutes the genome of many important cultivars; for example, the plantain subgroup—a staple food for millions of people in West and Central Africa and Latin America–Caribbean regions (De Langhe et al. 2009; Jones 2000; Perrier et al. 2009; Simmonds 1962). Thus, the use of the B genome in crop improvement seems logical and essential to produce high-yielding, disease-resistant cultivars.

An important factor limiting banana production is the prevalence of pests and diseases. Banana streak disease, first reported in Ivory Coast in 1966 (Lassoudière 1974; Yot-Dauthy and Bové 1966), is caused by a complex of BSV (Gayral and Iskra-Caruana 2009; Geering et al. 2000; Harper et al. 2005). All BSV are transmitted in a semipersistent manner by at least one of three species of mealybug, of which *Planococcus citri* is the most prevalent (Kubiriba et al. 2001; Lheureux 2002; Lockhart and Jones 2000; Meyer et al. 2008). A schematic representation of the triangular pathogenic BSV–banana–mealybug interaction is illustrated in Figure 1A. Propagation of the disease is known to be slow and easy to control. Field observations suggest that natural dissemination of BSV by mealybugs is limited and does not play a major role in epidemiology. The most significant transmission is likely to occur through vegetative propagation by exchanges of suckers and plantlets (Lockhart and Jones 2000). BSV infections cause characteristic discontinuous chlorotic areas turning into necrotic streaks on leaves, and a split pseudostem (Fig. 1A); a severe form develops heart rot of the pseudostem, which eventually leads to death of the plants (Lockhart and Jones 2000). However, a broad range of symptom expression, including symptomless plants, is reported depending upon several factors, including virus species, host genotype, level of management, and environmental conditions such as temperature and water stress (Dahal et al. 1998b; Gauhl and Pasberg-Gauhl 1995; Lockhart and Jones 2000; Lockhart 1986), making diagnosis of the disease difficult. Finally, various estimates of yield loss due to

BSV infection ranging from 6 to 90% are described in the literature (Dahal et al. 2000; Daniells et al. 2001; Lassoudière 1974). Today, epidemics are restricted to Africa, mainly Uganda and Nigeria (Kubiriba et al. 2001; Harper et al. 2004, 2005).

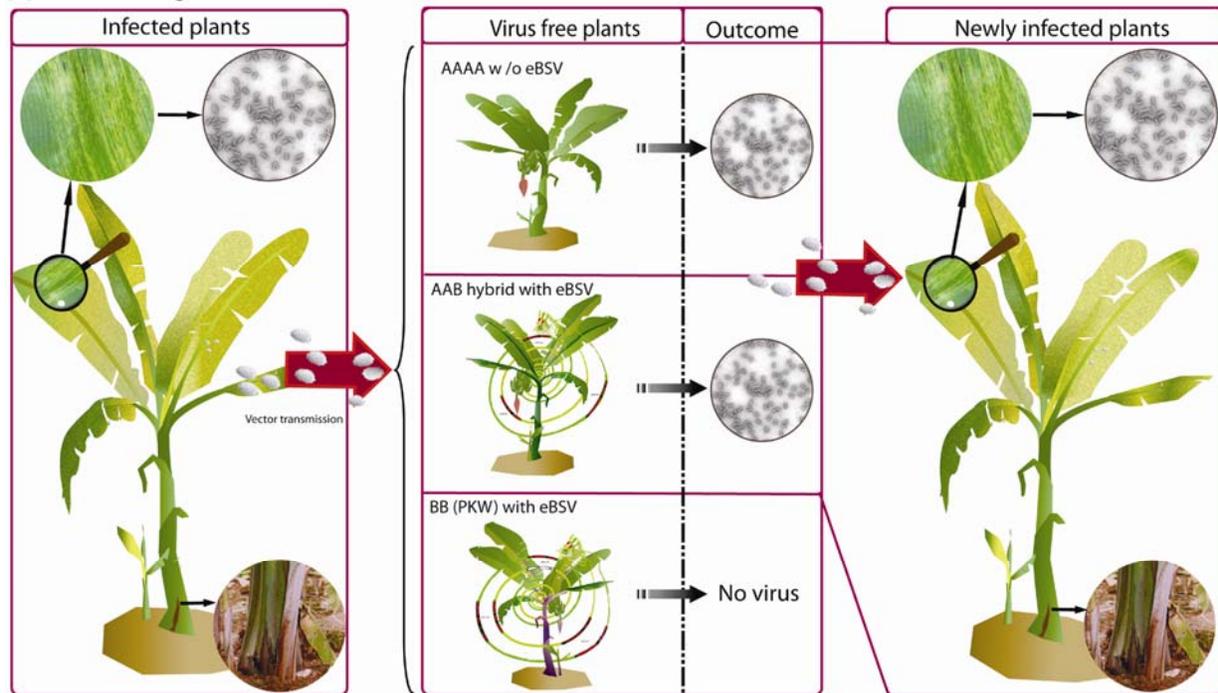
BSV.

BSV are nonenveloped plant pararetroviruses belonging to the genus *Badnavirus* within the family *Caulimoviridae* (Fauquet et al. 2005; Lockhart 1986). They have bacilliform-shaped virions (Fig. 1A), 30 by 150 nm in size, and a double-stranded, noncovalently linked circular DNA genome of 7.4 kb, which is replicated by reverse transcription. Their genomes contain three open reading frames (ORF) on one strand. ORF I and II potentially encode two small proteins of unknown function of 20.8 and 14.5 kDa, respectively. ORF III encodes a polyprotein of 208 kDa consisting of a putative cell-to-cell movement protein, the coat protein, an aspartic protease, and the viral replicase, which has reverse transcriptase (RT) and ribonuclease H (RNase H) functions (Harper and Hull 1998). This latter polyprotein is thought to be post-translationally cleaved into functional units by the aspartic protease. BSV have a wide serological and molecular variability. Indeed, BSV is considered to be the generic name of several distinct virus species, with similar biological properties, all infecting banana. Studies of partial sequences of BSV collected in Uganda and Mauritius (Geering et al. 2005; Harper et al. 2004; Jaufferally-Fakim et al. 2006) revealed great genetic diversity, with up to 30% nucleotide divergence among BSV isolates infecting the same *Musa* host plant. According to the International Committee on Taxonomy of Viruses, *Badnavirus* spp. are defined when viruses differ in their nucleotide sequence by more than 20% in the polymerase (RT + RNase H) region (Fauquet et al. 2005). Phylogenetic analyses of *Badnavirus* species, including BSV sequences belonging to either episomal viruses or endogenous viral sequences, confirmed the polyphyletic organization of BSV sequences into three distinct groups (Gayral and Iskra-Caruana 2009) as previously described by Harper and associates (2005). Several viral sequences from BSV isolates found among groups 2 and 3 were closely related to viruses that do not naturally infect banana, such as *Sugarcane bacilliform Mor virus*, *Commelina yellow mosaic virus*, *Cacao swollen shoot virus*, and *Citrus mosaic badna virus*, with viruses more distantly related to known BSV isolates (e.g., BSOLV, BSI_mV, BSGFV, *Banana streak Cav virus*, *Banana streak acuminata VN virus*, and *Banana streak Mys virus* forming group 1 and infecting banana). Thus, the information now available on BSV genetics indicates that a revision of current *Badnavirus* taxonomy is required (Gayral and Iskra-Caruana 2009).

Discovery of endogenous BSV.

Although epidemics are restricted to east-African countries and Nigeria, the last 20 years have seen BSV become one of the most widely distributed viruses of banana (Fargette et al. 2006; Lockhart and Jones 2000). However, BSV was not considered a serious threat until recently, when numerous spontaneous outbreaks of the disease occurred among micropropagated interspecific *Musa* hybrids and promising banana breeding lines in all banana-producing areas with no external viral source as the origin. The use of molecular BSV detection tools such as direct polymerase chain reaction (PCR) indicated the widespread presence of viral sequences in *Musa* plants (LaFleur et al. 1996); however, disease symptoms or viral particles were not systematically associated. The origin of these field outbreaks was correlated with the presence of infectious endogenous BSV (eBSV) sequences present within the *M. balbisiana* genome (Côte et al. 2010; Dallot et al. 2001; Harper et

A Classic triangular BSV interaction



B Special quartet BSV interaction

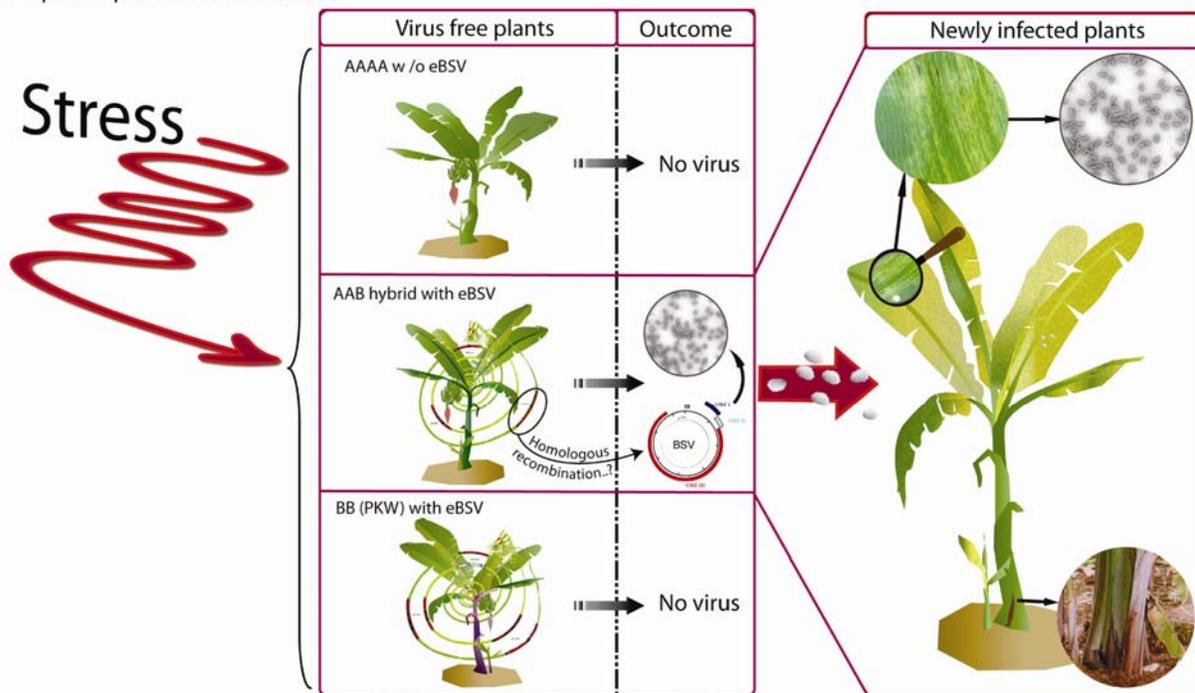


Fig. 1. Schematic representation of the two means of *Banana streak virus* (BSV) spread through either **A**, a classic triangular interaction or **B**, an unusual quartet interaction. **A**, In the classic triangular interaction, BSVs are vector transmitted (mealybugs) from infected plant (primary inoculum source) to virus-free plants, which then become infected and available (secondary inoculum source) to spread disease. BSV infections cause characteristic discontinuous chlorotic areas turning into necrotic streaks on leaves, and a split pseudostem. BSV has bacilliform-shaped virions. BSV spread is illustrated for three kinds of virus-free plants with different genotypes: a plant without endogenous (e)BSV (AAAA genome); an interspecific hybrid plant (AAB) harboring inducible infectious eBSV (only one copy of the B genome); and a wild *Musa balbisiana* diploid plant (Pisang Klutuk Wulung [PKW]), which is a carrier of infectious eBSV. Disease development occurs in both plants without (w/o) eBSV and interspecific hybrid plants harboring eBSV. Those plants then become virus reservoirs for further horizontal or vector transmission. In contrast, BSV infection never occurs in PKW via vector transmission; PKW appears resistant to BSV. The mechanism underlying this resistance is still unknown. **B**, In the unusual quartet BSV interaction, primary infection is due to BSV arising from within via eBSV. Of the three plant genotypes described in A, only two harbor eBSV. Following stress triggering eBSV expression, only interspecific hybrid plants are able to be infected. The mechanism underlying the transformation from integrated form to episomal viral genome is still uncharacterized in the case of BSV, although homologous recombination is strongly suspected. This hybrid then becomes a reservoir for further horizontal or vector transmission, as described in A. The diploid BB plant PKW, which harbors the same infectious integrations as hybrid plants but at a mono- or di-allelic level, presents a natural resistance to this internal activation. Here again, the mechanism underlying this resistance is still unclear, although virus resistance could occur, as in petunia and tobacco (Mette et al. 2002; Noreen et al. 2007), via homology-dependent transcriptional or posttranscriptional gene silencing.

al. 1999; Ndowora et al. 1999; Lheureux et al. 2003). Different abiotic stresses have been identified that trigger the production of viruses from infectious eBSV: temperature differences, water stress, in vitro culture, and interspecific crosses (Côte et al. 2010; Dahal et al. 1998b, 2000; Dallot et al. 2001; Lheureux et al. 2003) (Fig. 1B). Indeed, extensive studies on micropropagation procedures have clearly shown that the proliferation stage was a major determinant in the spontaneous appearance of viral particles BSV in interspecific *Musa* cultivars, regardless of the nature of the hybrids used (e.g., whether synthetic [AAAB and AAB] or natural, such as plantains [AAB]) (Côte et al. 2010; Dallot et al. 2001). This indicates that such activation is a general phenomenon. Nevertheless, little is known about the exact mechanisms underlying the expression of functional viral genomes from eBSV.

The second context involved in triggering virus production from infectious eBSV is genetic hybridization (Lheureux et al. 2003). A surprisingly high incidence of BSV was recorded among an F1 triploid (AAB) population obtained from interspecific hybridization between episomal virus-free diploid *M. balbisiana* (BB) and tetraploid *M. acuminata* (AAAA) parents, without there being any possibility of contamination. BSV infection regularly appears in one-half of the F1 progeny, indicating that genetic mechanisms are involved. This was characterized as a monogenic allelic system depending on a genetic factor called BEL for BSV Expressed Locus (Lheureux et al. 2003) conferring the role of carrier on the *M. balbisiana* diploid parent. Further studies revealed that at least one other factor (i.e., the ploidy of the B genome in *Musa* genotypes) is involved in BSV expression. Interestingly, and surprisingly, diploid *M. balbisiana* such as Pisang Klutuk Wulung (PKW) and Pisang Batu used as

female parents harbor infectious eBSV but are nevertheless resistant to BSV (both episomal and integrated forms).

These results have profound implications for the management of eBSV in interspecific *Musa* cultivars. So far, it has not been possible to identify which banana plants harboring integrated BSV sequences will release virions when submitted to a stress. However, it is now obvious that outbreaks of the disease and BSV epidemics arise as a consequence of both the activation of eBSV as well as mealybug spread of exogenous forms of the virus. Today, banana streak disease due to eBSV is an important constraint to banana germplasm movement, genetic improvement efforts, and mass propagation.

Banana streak disease—a remarkable pathogenic quartet interaction.

In light of the observations and results described above, the banana–BSV–mealybug pathosystem can, in fact, be viewed as a quartet involving eBSV as an alternative partner to release virions (Fig. 1B).

The genome of *M. balbisiana* contains eBSV of several species, and three natural widespread species have so far been identified as infectious: BSOLV, BSI_mV, and BSGFV (Gayral et al. 2008; Geering et al. 2005; Harper et al. 1999; Iskra-Caruana et al. 2003; Ndowora et al. 1999). Although infectious, their presence alone is not enough to induce infection (Gayral et al. 2008). Indeed, it appears that other levels of regulation (probably genetic; described in the section on eBSV discovery) are required to release a virus particle from its integrated counterpart. These different eBSV could be activated in the same banana plant regardless of the stress applied (Côte et al. 2010; Lheureux et al. 2003).

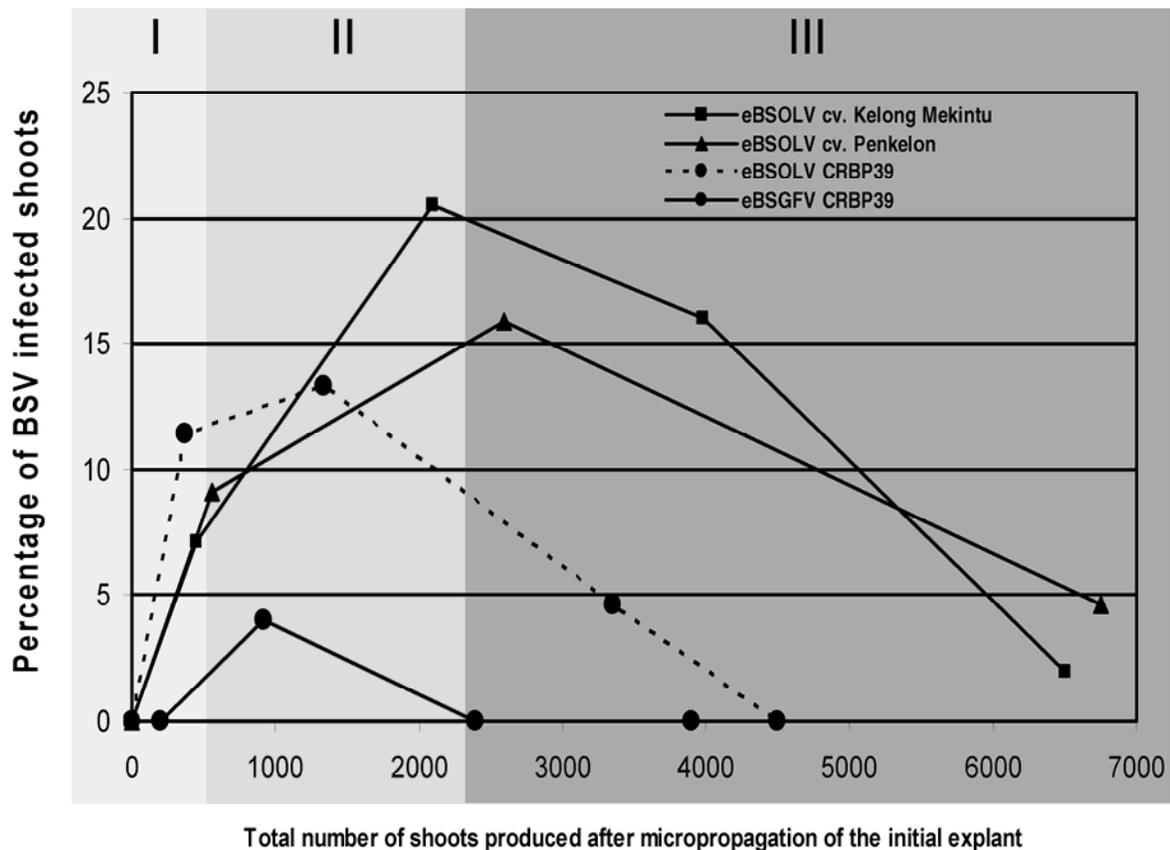


Fig. 2. Activation patterns of infectious endogenous *Banana streak virus* (eBSV) in two natural banana cultivars (Kelong Mekintu and Penkelon) and a synthetic banana hybrid (CRBP39). Curves represent the percentage of shoots infected with *Banana streak OL virus* (BSOLV) in Kelong Mekintu (■), in Penkelon (▲), and in CRBP 39 (●) and the percentage of shoots infected with *Banana streak GF virus* (BSGFV) in CRBP 39 (● on dotted line). Numbers I, II, and III represent the three distinct phases of eBSV activation.

Dallot and associates (2001) demonstrated that shoot tip micropropagation during in vitro culture triggers the expression of infectious eBSOLV in the synthetic tetraploid interspecific hybrid cv. FHIA21 (AAAB), and Côte and associates (2010) showed that the same was true for eBSOLV and eBSGFV in cv. CRBP39, another synthetic tetraploid interspecific hybrid (AAAB), as well as in cvs. Kelong Mekintu and Black Penkelon, two natural triploid plantains (AAB). During such in vitro culture experiments, the percentage of infected plantlets was recorded. Similar curves of infected plantlets were observed for all cultivars studied (Fig. 2). Three distinct phases were observed: i) a strong increase during the initial subculture, ii) a slower increase until the highest percentage of infected plants was reached and immediately afterward, and iii) a decrease that could go down to zero according to the cultivar studied. The maximum percentages of infected plantlets ranged from 5 to 20%, depending on the cultivar and eBSV analyzed, but values were consistently below 30%. The resulting curve could arise from the conjunction of three distinct biological phenomena: i) the activation of infectious eBSV, leading to a consequent increase in the number of BSV-infected plants; ii) a concomitant decrease in the episomal particle titer in neofomed plantlets due to cell multiplication outcompeting virus replication; and iii) activation of infectious eBSV in a limited number of cells that did not occur throughout the entire shoot tip (Côte et al. 2010).

Among the various cultivars tested, cv. CRBP39 was the only one to display activation patterns of both infectious eBSOLV and eBSGFV. This could result from differences in the regulation of eBSGFV expression because all cultivars tested harbored the same pattern of integration (i.e., same integration loci in the *Musa* genome and the same eBSV molecular structure) (Côte et al. 2010). Infectious eBSOLV and eBSGFV also displayed different maximum activation levels in cv. CRBP39, indicating that the activated cells could express BSOLV, BSGFV, or both BSOLV and BSGFV, and that independent regulation of expression of each infectious eBSV exists in cv. CRBP39. All these results indicate that the activation potential for a given infectious eBSV differs between cultivars and between various infectious eBSV within given interspecific *Musa* genotypes.

In parallel, the genomic structure of eBSV in banana was investigated in the wild diploid *M. balbisiana* PKW (BB). A bacterial artificial chromosome (BAC) library was constructed (Safar et al. 2004) and explored for the pattern of integration of eBSV by testing with different sets of viral probes individually representing the complete genomes of BSOLV, BSGFV, and BSImV. A very small number of hits were recorded for each BSV species. The BAC identified differed from one BSV species to another, indicating independent low-copy integration. BSV-positive BAC clones were characterized by restriction fragment length polymorphism (RFLP) fingerprinting approaches. To date, we have demonstrated that PKW harbors at least five different BSV integrants: two corresponding to BSOLV, two to BSGFV, and one to BSImV (F. C. Baurens, M. Chabannes, S. Bocs, V. Barbe, P. Gayral, P. O. Debroy, A. D'Hont, and M.-L. Iskra-Caruana, *unpublished data*). One BAC clone for each integrant has been sequenced and fully annotated. All integrants—for example, eBSGFV (named eBSGFV-7 and eBSGFV-9), eBSImV, and eBSOLV (named eBSOLV-1 and eBSOLV-2)—have been fully characterized (Gayral et al. 2008, 2010). In all cases, the integrants are much longer than a single BSV genome (7.26 kb): 13.28 and 15.58 kb for eBSGFV, 15.8 kb for eBSImV, and 22.9 kb and 23.2 kb for eBSOLV. All exhibit a more or less complex rearrangement of viral sequences, with most viral regions being duplicated

and, therefore, present in several copies within each eBSV, either in the same or opposite orientation with respect to the organization of the episomal BSV genome. For the sake of simplicity, we will focus mainly on BSGFV integrations. Both eBSGFVs comprised six juxtaposed viral fragments (Fig. 3). These six fragments are closely related to each other because they share five common regions (I to IV and VI) (Fig. 3) with 99.8% nucleotide sequence identity; eBSGFV-9 carries an additional segment on fragment V. They share a very high degree of similarity (99.3% identity on average) with episomal BSGFV particles. Moreover, eBSGFV-9 has accumulated 35 more mutations than eBSGFV-7, 3 of them leading to null mutations, which is discriminating in terms of virus genome functionality (Gayral et al. 2008). eBSImV appeared less rearranged and complex than eBSGFV and eBSOLV (F. C. Baurens, M. Chabannes, S. Bocs, V. Barbe, P. Gayral, P. O. Debroy, A. D'Hont, and M.-L. Iskra-Caruana, *unpublished data*), despite being composed of five fragments (Gayral et al. 2010). Fragments I and II are very short (30 and 27 nucleotides), while fragment III (13.7 kbp) contains 170% of the BSImV genome integrated in one direction.

Although the viral genome appears fragmented, inverted, and partially repeated in each eBSV, surprisingly, at least one integrant for each BSV species contains the full-length genome of BSV at least once (Gayral et al. 2008) (F. C. Baurens, M. Chabannes, S. Bocs, V. Barbe, P. Gayral, P. O. Debroy, A. D'Hont, and M.-L. Iskra-Caruana, *unpublished data*).

Overall, a high degree of sequence conservation between the integrated and episomal form of the virus can be observed, suggesting either a recent integration event or the recent emergence of these BSV species. Genotyping molecular markers (PCR and PCR-RFLP) were developed to distinguish both the two integrant types and their locus of integration in the banana chromosome. eBSV segregation among the AAB F1 progeny from the interspecific genetic cross involving PKW (BB) as the female parent and cv. IDN 110T (AAAA) as the male parent was analyzed. Each integrated BSV species is allelic, which means that the integrated sequences of BSGFV, BSOLV, and BSImV in PKW result from a single integration event (Gayral et al. 2008, 2010). The integration of BSGFV and BSOLV appears di-allelic, whereas that of BSImV is mono-allelic. Finally, in the case of BSGFV, Gayral and associates (2008) have demonstrated that only one allele (eBSGFV-7) is infectious.

Taking all the available results together, including activation in in vitro culture and the genetic segregation of episomal viruses in the resulting progeny, provides evidence of the genomic origin of disease caused by a virus that does not require genomic integration to replicate. However, the mechanism that allows formation of a circular viral genome from its integrated counterpart remains unknown. Based on analysis of the eBSGFV structure, and because the direct transcription event that happens in petunia (Richert-Poggeler et al. 2003; Noreen et al. 2007) seems unlikely to occur in our situation, we propose a theoretical activation model (Fig. 4) based on homologous recombination (HR). This model is based on a minimum number of HR steps to reconstitute a full-length circular BSGFV genome from eBSGFV-7. In this model, a first recombination, called Rec1, occurs between two 642-bp inverted repeats separated by 6,942 kbp (Fig. 4A). This intramolecular recombination results in inversion of an internal sequence and leads to a juxtaposition of fragments II and V (Fig. 4B). The intramolecular reorganization of eBSV generates a full-length linear BSGFV genome containing the three ORF in the same orientation (Fig. 4B).

This viral genome is flanked by two 633-bp direct repeats, which are involved in the final HR step. HR Rec2 produces an

eBSV with a deletion of 7,324 bp called “modified eBSGFV-7 locus” and, more interestingly, results in the excision (“pop-out”) of a circular viral molecule resembling the BSGFV circular genome (Fig. 4C).

According to this model, activation of eBSGFV-7 could be realized with only two steps of HR. In this scheme, only the two fragments II and V contributed in fine to eBSV activation. The in silico analysis of the completeness of the fragments involved in the reconstitution of the BSGFV genome is compatible with our model. Indeed, the sequence of fragments II and V in eBSGFV-7 is free of any mutations causing a loss of viral function, which is not the case for eBSGFV-9 (Fig. 3) (Gayral et al. 2008). This model corroborates data presented by Gayral and associates (2008), who showed that only the eBSGFV-7 allele was infectious.

Consequences of the presence of a fourth partner.

The banana genome appears permeable to illegitimate BSV integrations although integration is not an essential step in the BSV replication cycle. To explain the presence and maintenance of eBSV in banana from an evolutionary point of view, several hypotheses have been invoked and tested, resulting in evidence that the fourth partner greatly impacts both the virus and the host plant.

Preliminary results have shown that eBSV are present in few copies per genome (Harper et al. 1999; Gayral et al. 2008, 2010), and mainly in *M. balbisiana* genomes (Gayral and Iskra-Caruana 2009). Large-scale phylogenetic studies comprising BSV and eBSV showed that integrations were frequent—27 independent integrations were recorded in banana—and recent, with most having occurred after *M. acuminata*/*M. balbisiana* speciation (Gayral and Iskra-Caruana 2009), estimated at 4.6 million years ago by Lescot and associates (2008). In contrast

to eBSV, ePVCV and eTVCV underwent large-scale genome amplification, probably mediated by retro elements found near EPRV, and have now reached hundreds to thousands of copies per host genome (Gregor et al. 2004; Mette et al. 2002; Richert-Poggeler et al. 2003; Staginnus et al. 2007). These differences from eBSV could result from the maintenance of ePVCV and eTVCV in their host genome for much longer periods of time. Most eBSV tend to accumulate random mutations, leading to defective ORF and, thus, noninfectious integrations (Gayral and Iskra-Caruana 2009). Although only a small fraction of eBSV is infectious, even this small fraction represents a significant partner in the banana–BSV pathosystem.

As previously discussed, by releasing functional viral genomes from which viral multiplication and plant infection can arise, infectious eBSV have a real impact on the outbreak of banana streak disease recorded in countries or continents previously free of BSV (Dahal et al. 1998a,b; Harper et al. 2004, 2005). Plants harboring infectious eBSV are sources of viruses and certainly contribute to the global biodiversity of recorded BSV. Therefore, the large molecular diversity reported so far for BSV results mainly from a mixture of episomal and integrated viral sequences. The recent studies on BSV phylogeny conducted by Gayral and Iskra-Caruana (2009) demonstrated that endogenous badnavirus sequences have certainly played a significant role in *Musa–Badnavirus* parasitic interactions, because few molecular differences occurred between infectious eBSV and their episomal viral counterparts (Gayral et al. 2008). Their regular activation consequently contributes to the diversity and evolution of today’s episomal BSV. Thus, eBSV can be considered as a reservoir of BSV biodiversity in that they contribute to the BSV gene pool once activated. A detailed analysis of the evolutionary history of infectious eBSImV and eBSGFV in relation to their banana host plant showed that their integra-

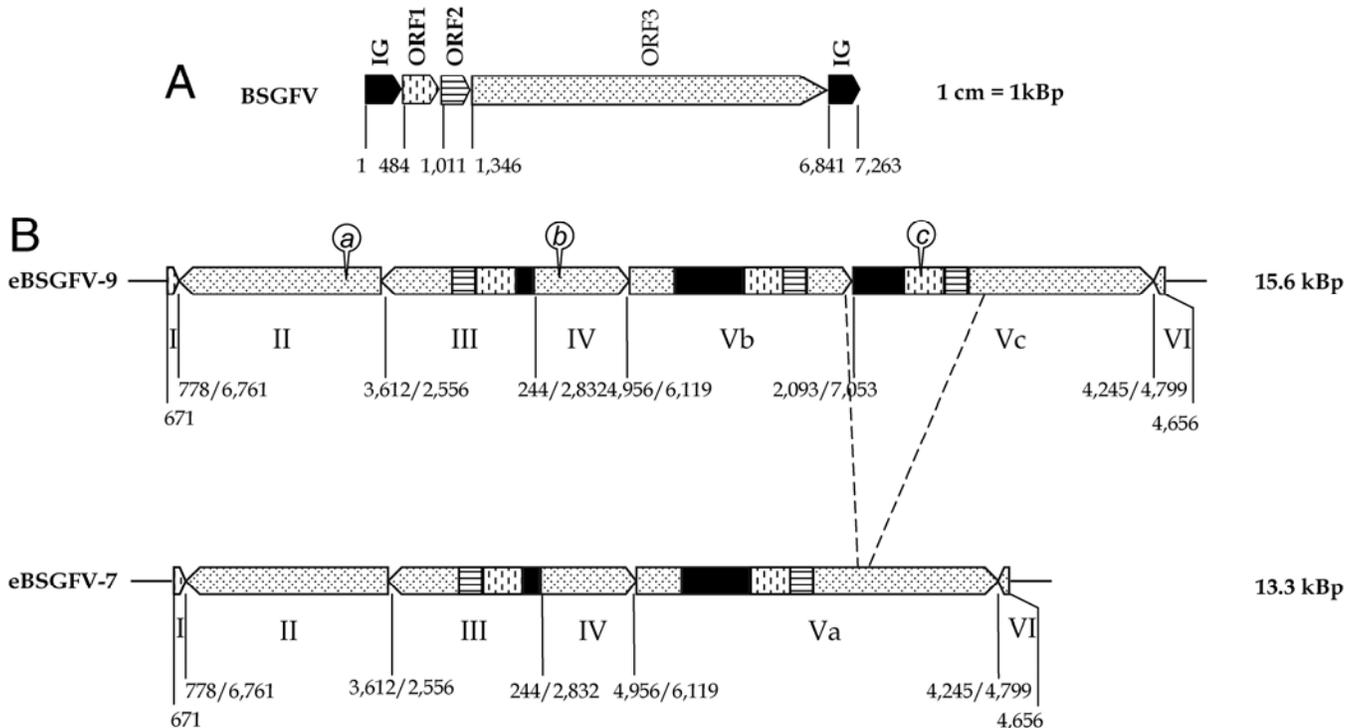


Fig 3. Endogenous *Banana streak GF virus* (eBSGFV) structures in Pisang Klutuk Wulung (adapted from Gayral et al. 2008). **A**, Organization (linear view) of the BSGFV genome (GenBank AY3509). Vertical hyphens, horizontal lines, and dotted lines indicate open reading frame (ORF)1, ORF2, and ORF3 of the virus, respectively. Intergenic region is shown in black. **B**, Structures of eBSGFV-9 (top) and eBSGFV-7 (bottom) resulting from annotated BAC clones MBP_94I16 and MBP_71C19, respectively. Letters a, b, and c above eBSGFV-9 represent null mutations in the viral sequence. Arrows indicate the orientation of fragments of the BSGFV genome integrated into the *Musa* genome (shown by thin horizontal lines). Different viral ORFs are indicated as in A. Roman numerals identify the viral fragments. Numbers below indicate the position of each fragment in the BSGFV genome.

tion was more recent than previously thought, circa 640,000 years ago (Gayral et al. 2010), and, in both cases, contributes to BSV diversity.

The hypothesis that eBSV sequences could confer virus resistance to the host plant may explain the observation that eBSV, and EPRV sequences in general, have been maintained in host genomes despite the fact that some integrants are harmful when activated. A low transcriptional level of EPRV, and sub-

sequent small-interfering RNA production was observed in petunia (*Petunia hybrida*) and tomato (*Solanum lycopersicum*) (Noreen et al. 2007; Staginnus et al. 2007). EPRV-induced homology-dependent gene silencing would target counterpart nonintegrated viruses: PVCV (genus *Petuvirus*) in petunia and TVCV (genus *Cavemovirus*) in tomato. This mechanism is also suspected for BSV in the wild diploid *M. balbiana* PKW, which is resistant to both activation of infectious eBSV

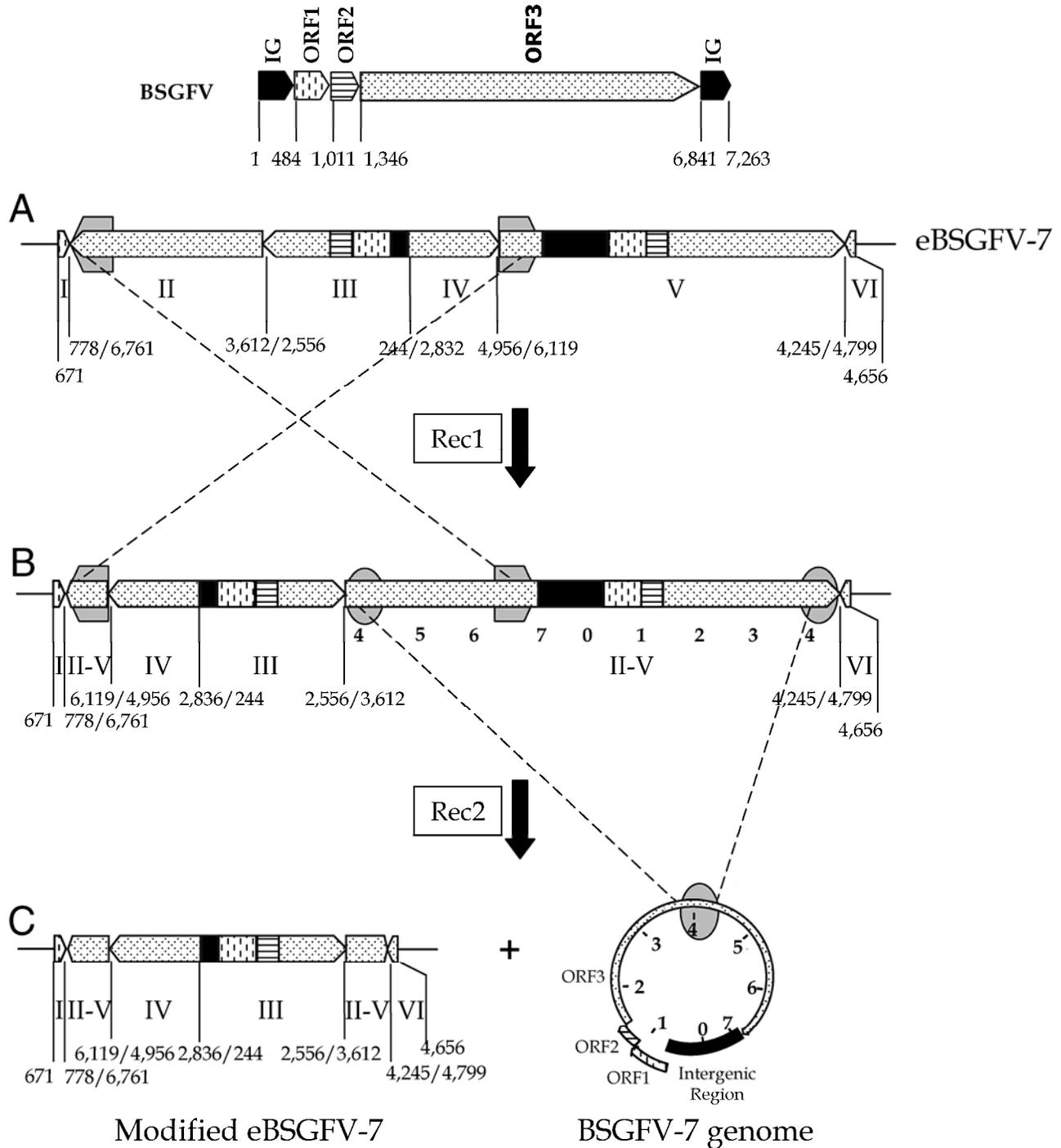


Fig. 4. Homologous recombination (HR)-based model to explain release of a *Banana streak GF virus* (BSGFV) genome from its endogenous counterpart, eBSGFV-7. Thin vertical lines delimit internal fragments of eBSGFV (I through VI) and numbers below vertical lines (Bp) and below fragments (numbered 1 to 7; kbp) refer to positions in the BSGFV genome (GenBank: AY493509). Thin horizontal lines represent flanking regions of eBSGFV in genomic *Musa* sequences. Repeats used as templates for each HR event are indicated by gray symbols. **A**, Schematic representation of native eBSGFV-7. **B**, Theoretical structure of the recombined molecule of eBSGFV-7 after the first recombination event (Rec1). **C**, Theoretical structures of recombined molecules after the second recombination event (Rec2). Left panel: residual eBSGFV remaining in the *Musa* genome. Right panel: circular full-length viral molecule excised from eBSGFV-7.

and BSV infection (Hohn et al. 2008; Lheureux et al. 2003). For instance, the infectious allele eBSGFV-7 exhibits a remarkable conservation through time and among *M. balbisiana* genotypes despite its potentially deleterious effects for interspecific hybrids, probably in line with a function brought by this integrant (Gayral et al. 2010).

Resistance mechanisms against episomal and integrated forms of members of the *Caulimoviridae* family are seen as a counter-adaptation of banana in response to the presence of infectious eBSV in the host genome. Using the same raw material, evolution has found a way to protect the host from the consequences of harmful integrations and from exogenous BSV, illustrating what François Jacob in 1977 called “evolution by molecular tinkering”.

ACKNOWLEDGMENTS

P. Gayral was supported by a Ph.D. grant, CIRAD-Région Languedoc Roussillon. We thank Serge Galzi for preparation of Figures 1 and 2.

LITERATURE CITED

- Ashby, M. K., Warry, A., Bejarano, E. R., Kashoggi, A., Burrell, M., and Lichtenstein, C. P. 1997. Analysis of multiple copies of geminiviral DNA in the genome of four closely related *Nicotiana* species suggest a unique integration event. *Plant Mol. Biol.* 35:313-321.
- Bakry, F., Carreel, F., Jenny, C., and Horry, J.-P. 2009. Genetic improvement of banana. Pages 3-50 in: *Breeding Plantation Three Crops: Tropical Species*. J. S. Mohan and P. M. Priyadarshan, eds. Academic Springer, New York.
- Bejarano, E., Khashoggi, A., Witty, A., and Lichtenstein, C. 1996. Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. *Proc. Natl. Acad. Sci. U.S.A.* 93:759-764.
- Budiman, M. A., Mao, L., Wood, T. C., and Wing, R. A. 2000. A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. *Genome Res.* 10:129-136.
- Côte, F.-X., Galzi, S., Folliot, M., Lamagnere, Y., Teycheney, P.-Y., and Iskra-Caruana, M.-L. 2010. Micropropagation by tissue culture triggers differential expression of endogenous *Banana streak virus* (eBSV) in the B genome of natural and synthetic interspecific banana plantain. *Mol. Plant Pathol.* 11:137-144.
- Dahal, G., D'Hugues, J. A., and Lockhart, B. E. L. 1998a. Status of banana streak disease in Africa: Problems and future research needs. *Integr. Pest Manage. Rev.* 3:85-97.
- Dahal, G., D'Hughes, J. A., Thottappilly, G., and Lockhart, B. E. L. 1998b. Effect of temperature on symptom expression and reliability of *Banana streak badnavirus* detection in naturally infected plantain and banana (*Musa* spp.). *Plant Dis.* 82:16-21.
- Dahal, G., Ortiz, R., Tenkouano, A., D'Hughes, J. A., Thottappilly, G., Vuylsteke, D., and Lockhart, B. E. L. 2000. Relationship between natural occurrence of *banana streak badnavirus* and symptom expression, relative concentration of viral antigen, and yield characteristics of some micropropagated *Musa* spp. *Plant Pathol.* 49:68-79.
- Dalot, S., Acuña, P., Rivera, C., Ramirez, P., Côte, F., Lockhart, B. E. L., and Caruana, M.-L. 2001. Evidence that the proliferation stage of micropropagation procedure is determinant in the expression of *Banana streak virus* integrated into the genome of the FHIA 21 hybrid (*Musa* AAAB). *Arch. Virol.* 146:2179-2190.
- Daniells, J. W., Geering, A. D. W., Bryde, N. J., and Thomas, J. E. 2001. The effect of *Banana streak virus* on the growth and yield of dessert bananas in tropical Australia. *Ann. Appl. Biol.* 139:51-60.
- De Langhe, E., Vrydaghs, L., De Maret, P., Perrier, X., and Denham, T. 2009. Why bananas matter: An introduction to the history of banana domestication. *Ethnobot. Res. Appl.* 7:165-177.
- Fargette, D., Konate, G., Fauquet, C., Muller, E., Peterschmitt, M., and Thresh, J. M. 2006. Molecular ecology and emergence of tropical plant viruses. *Annu. Rev. Phytopathol.* 44:235-260.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., and Ball, L. A. 2005. *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*. Academic Press, London.
- Gambley, C. F., Geering, A. D. W., Steele, V., and Thomas, J. E. 2008. Identification of viral and non-viral reverse transcribing elements in pineapple (*Ananas comosus*), including members of two new *Badnavirus* species. *Arch. Virol.* 153:1599-1604.
- Gauhl, F., and Pasberg-Gauhl, C. 1995. Temporal dynamics of *banana streak virus* (BSV) symptoms in *Musa* clones in southeast Nigeria. *Phytopathol.* 25:27.
- Gayral, P., and Iskra-Caruana, M.-L. 2009. Phylogeny of *Banana streak virus* reveals recent and repetitive endogenization in the genome of its banana host (*Musa* sp.). *J. Mol. Evol.* 69:65-80.
- Gayral, P., Noa, C., Lescot, M., Lheureux, F., Lockhart, B., Matsumoto, T., Piffanelli, P., and Iskra-Caruana, M.-L. 2008. A single *banana streak virus* integration event in the banana genome as the origin of infectious endogenous pararetrovirus. *J. Virol.* 82:6697-6710.
- Gayral, P., Blondin, L., Guidolin, O., Carreel, F., Hippolyte, I., Perrier, X., and Iskra-Caruana, M.-L. 2010. Evolution of endogenous sequences of *Banana streak virus*: What can we learn from banana (*Musa* sp.) evolution? *J. Virol.* 84:7346-7359.
- Geering, A. D. W., McMichael, L. A., Dietzgen, R. G., and Thomas, J. E. 2000. Genetic diversity among *Banana streak virus* isolates from Australia. *Phytopathology* 90:921-927.
- Geering, A. D. W., Olszewski, N., Dahal, G., Thomas, J. E., and Lockhart, B. E. L. 2001. Analysis of the distribution and structure of integrated *Banana streak virus* DNA in a range of *Musa* cultivars. *Mol. Plant Pathol.* 2:207-213.
- Geering, A. D. W., Olszewski, N. E., Harper, G., Lockhart, B. E. L., Hull, R., and Thomas, J. E. 2005. Banana contains a diverse array of endogenous badnaviruses. *J. Gen. Virol.* 86:511-520.
- Gregor, W., Mette, M. F., Staginnus, C., Matzke, M. A., and Matzke, A. J. M. 2004. A distinct endogenous pararetrovirus family in *Nicotiana tomentosiformis*, a diploid progenitor of polyploid tobacco. *Plant Physiol.* 134:1191-1199.
- Hansen, C. N., Harper, G., and Heslop-Harrison, J. S. 2005. Characterisation of pararetrovirus-like sequences in the genome of potato (*Solanum tuberosum*). *Cytogenet. Genome Res.* 110:559-565.
- Harper, G., and Hull, R. 1998. Cloning and sequence analysis of *banana streak virus* DNA. *Virus Genes* 17:271-278.
- Harper, G., Osuji, J. O., Heslop-Harrison, P., Hull, R., Harper, G., and Lockhart, B. E. L. 1999. Integration of *Banana streak badnavirus* into the *Musa* genome: Molecular evidence. *Virology* 255:207-213.
- Harper, G., Hull, R., Lockhart, B. E. L., and Olszewski, N. 2002. Viral sequences integrated into plant genomes. *Annu. Rev. Phytopathol.* 40:119-136.
- Harper, G., Richert-Poggeler, K. R., Hohn, T., and Hull, R. 2003. Detection of *Petunia vein-clearing virus*: Model for the detection of DNA viruses in plants with homologous endogenous pararetrovirus sequences. *J. Virol. Methods* 107:177-184.
- Harper, G., Hart, D., Moul, S., and Hull, R. 2004. *Banana streak virus* is very diverse in Uganda. *Virus Res.* 100:51-56.
- Harper, G., Hart, D., Moul, S., Hull, R., Geering, A., and Thomas, J. 2005. The diversity of *Banana streak virus* in Uganda. *Arch. Virol.* 150:2407-2420.
- Heslop-Harrison, J. S., and Schwarzacher, T. 2007. Domestication, genomics and the future for banana. *Ann. Bot.* 100:1073-1084.
- Hohn, T., Richert-Poggeler, K. R., Harper, G., Schwarzacher, T., Teo, C., Teycheney, P.-Y., Iskra-Caruana, M.-L., and Hull, R. 2008. Evolution of integrated plant viruses. Pages 54-76 in: *Plant Virus Evolution*. M. Roosinck, ed. Academic Press, Heidelberg, Germany.
- Hull, R. 2001. Classifying reverse transcribing elements: A proposal and challenge to the ICTV. *Arch. Virol.* 146:2255-2261.
- Hull, R. 2002. *Matthew's Plant Virology*. Academic Press, London.
- Hull, R., Harper, G., and Lockhart, B. E. L. 2000. Viral sequences integrated into plant genome. *Trends Plant Sci.* 5:362-365.
- Iskra-Caruana, M. L., Lheureux, F., Noa-Carrazana, J. C., Piffanelli, P., Carreel, F., Jenny, C., Laboureaux, N., and Lockhart, B. E. L. 2003. Unstable balance of relation between pararetrovirus and its host plant: The BSV-EPRV banana pathosystem. Page 8 in: *EMBO Workshop Genomic Approaches in Plant Virology*. Keszthely, Hungary.
- Jakowitsch, J., Mette, M. F., van der Winden, J., and Matzke, M. M. A. 1999. Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. *Proc. Natl. Acad. Sci. U.S.A.* 96:13241-13246.
- Jaufeerally-Fakim, Y., Khorughdarry, A., and Harper, G. 2006. Genetic variants of *Banana streak virus* in Mauritius. *Virus Res.* 115:91-98.
- Jones, D. R. 2000. *Diseases of Banana, Abaca and Enset*. CABI, Oxon, U.K.
- Kenton, A., Kashoggi, A., Parokony, A., Bennett, M. D., and Lichtenstein, C. 1995. Chromosomal location of endogenous geminivirus-related DNA sequences in *Nicotiana tabacum* L. *Chromosome Res.* 3:346-350.
- Kubiriba, J., Legg, J. P., Tushemereirwe, W., and Adipala, E. 2001. Vector transmission of *Banana streak virus* in the greenhouse in Uganda. *Ann. Appl. Biol.* 139:37-43.
- Kunii, M., Kanda, M., Nagano, H., Uyeda, I., Kishima, Y., and Sano, Y. 2004. Reconstruction of putative DNA virus from endogenous *rice tungro bacilliform virus*-like sequences in the rice genome: Implications for integration and evolution. *BMC Genomics* 5:80.

- LaFleur, D. A., Lockhart, B. E. L., and Olszewski, N. E. 1996. Portions of the banana streak badnavirus genome are integrated in the genome of its host *Musa* sp. (Abstr.) *Phytopathology* 86:S100.
- Lassoudière, A. 1974. La mosaïque dite "à tirets" du bananier Poyo en Côte d'Ivoire. *Fruits* 29:349-357.
- Lescot, M., Piffanelli, P., Ciampi, A. Y., Ruiz, M., Blanc, G., Leebens-Mack, J., Da Silva, F. R., Santos, C. M., D'Hont, A., Garsmeur, O., Vilarinhos, A. D., Kanamori, H., Matsumoto, T., Ronning, C. M., Cheung, F., Haas, B. J., Althoff, R., Arbogast, T., Hine, E., Pappas, G. J., Sasaki, T., Souza, M. T., Miller, R. N., Glaszmann, J.-C., and Town, C. D. 2008. Insights into the *Musa* genome: Syntenic relationships to rice and between *Musa* species. *BMC Genomics* 9:58.
- Lheureux, F. 2002. Les mécanismes génétiques d'activation et de multiplication du Banana streak virus intégré dans le génome *Musa* sp. Thesis. Ecole Nationale Supérieure Agronomique de Montpellier. ED Biologie Intégrative, Université Sciences et Techniques du Languedoc, France.
- Lheureux, F., Carreel, F., Jenny, C., Lockhart, B. E. L., and Iskra-Caruana, M.-L. 2003. Identification of genetic markers linked to Banana streak disease expression in inter-specific *Musa* hybrids. *Theor. Appl. Genet.* 106:594-598.
- Liu, R., Koyanagi, K., and Kishima, Y. 2009. The integration and rearrangement of the pararetrovirus in the rice genomes. In: 6th Int. Rice Genomics Symp. Manila, Philippines.
- Lockhart, B. E. L. 1986. Purification and serology of a bacilliform virus associated with banana streak disease. *Phytopathology* 76:995-999.
- Lockhart, B. E. L., and Jones, D. 2000. Banana streak. Pages 263-274 in: *Diseases of Banana, Abaca and Enset*. D. Jones, ed. CAB International, Wallingford, U.K.
- Lockhart, B. E. L., Menke, G., Dahal, G., and Olszewski, N. E. 2000. Characterisation and genomic analysis of *Tobacco vein clearing virus*, plant pararetrovirus that is transmitted vertically and related to sequences integrated in the host genome. *J. Gen. Virol.* 81:1579-1587.
- Mao, L., Wood, T. C., Yu, Y., Budiman, M. A., Tomkins, J., Woo, S., Sasinowski, M., Presting, G., Frisch, D., Goff, S., Dean, R. A., and Wing, R. A. 2000. Rice transposable elements: A survey of 73 000 sequence-tagged-connectors. *Genome Res.* 10:982-990.
- Mette, M. F., Kanno, T., Aufsatz, W., Jakowitsch, J., van der Winden, J., Matzke, M. A., and Matzke, A. J. M. 2002. Endogenous viral sequences and their potential contribution to heritable virus resistance in plants. *EMBO (Eur. Mol. Biol. Organ.) J.* 21:461-469.
- Meyer, J. B., Kasdorf, G. G. F., Nel, L. H., and Pietersen, G. 2008. Transmission of activated-episomal *Banana streak OL* badnavirus (BSOLV) to cv. Williams banana (*Musa* sp.) by three mealybug species. *Plant Dis.* 92:1158-1163.
- Ndowora, T., Dahal, G., LaFleur, D., Harper, G., Hull, R., Olzsewski, N., and Lockhart, B. E. L. 1999. Evidence that badnavirus infection in *Musa* can originate from integrated pararetroviral sequences. *Virology* 255:214-220.
- Noreen, F., Akbergenov, R., Hohn, T., and Richert-Poggeler, K. R. 2007. Distinct expression of endogenous *Petunia vein clearing virus* and the DNA transposon dTph1 in two *Petunia hybrida* lines is correlated with differences in histone modification and siRNA production. *Plant J.* 50:219-229.
- Pahalawatta, V., Druffel, K., and Pappu, H. 2008. A new and distinct species in the genus *Caulimovirus* exists as an endogenous plant pararetroviral sequence in its host, *Dahlia variabilis*. *Virology* 376:253-257.
- Perrier, X., Bakry, F., Carreel, F., Jenny, C., Horry, J.-P., Lebot, V., and Hippolyte, I. 2009. Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot. Res. Appl.* 7:199-216.
- Richert-Poggeler, K. R., and Shepherd, R. J. 1997. *Petunia vein-clearing virus*: A plant pararetrovirus with the core sequences for an integrase function. *Virology* 236:137-146.
- Richert-Poggeler, K. R., Noreen, F., Schwarzacher, T., Harper, G., and Hohn, T. 2003. Induction of infectious *Petunia vein clearing (pararetro) virus* from endogenous provirus in petunia. *EMBO (Eur. Mol. Biol. Organ.) J.* 22:4836-4845.
- Safar, J., Noa-Carrazana, J. C., Vrana, J., Bartos, J., Alkhimova, O., Lheureux, F., Simkova, H., Iskra-Caruana, M.-L., Dolezel, J., and Piffanelli, P. 2004. Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. *Genome* 47:1182-1191.
- Simmonds, N. W. 1962. *The Evolution of the Bananas*. Longmans, Green and Co., Ltd., London.
- Simmonds, N. W., and Shepherd, K. 1955. The taxonomy and origins of the cultivated bananas. *J. Linn. Soc. Bot.* LV:302-312.
- Staginnus, C., and Richert-Poggeler, K. R. 2006. Endogenous pararetroviruses: Two-faced travelers in the plant genome. *Trends Plant Sci.* 11:485-491.
- Staginnus, C., Gregor, W., Mette, M. F., Teo, C. H., Borroto-Fernandez, E. G., Machado, M. L., Matzke, M., and Schwarzacher, T. S. 2007. Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. *BMC Plant Biol.* 7:24.
- Staginnus, C., Iskra-Caruana, M.-L., Lockhart, B. E. L., Hohn, T., and Richert-Poggeler, K. R. 2009. Suggestions for nomenclature of endogenous pararetroviral (EPRV) sequences in plants. *Arch. Virol.* 154:1189-1193.
- Yot-Dauthy, D., and Bové, J.-M. 1966. Mosaïque du bananier. Identification et purification de diverses souches du virus. *Fruits* 21:449-465.