

The genomes of several plant species contain endogenous geminiviral sequences

Denis Filloux, Sasha Murrell, Maneerat Koohapitagtam, Michael Golden, Charlotte Julian, Serge Galzi, Marilyne Uzest, Marguerite Rodier-Goud, Angélique d'Hont, Marie-Stéphanie Vernerey, et al.

▶ To cite this version:

Denis Filloux, Sasha Murrell, Maneerat Koohapitagtam, Michael Golden, Charlotte Julian, et al.. The genomes of several plant species contain endogenous geminiviral sequences. 8. International Geminivirus Symposium, Nov 2016, New Delhi, India. 1 p., 2016. hal-02798133

HAL Id: hal-02798133 https://hal.inrae.fr/hal-02798133v1

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The genomes of several plant species contain endogenous geminiviral sequences





<u>Denis Filloux</u>,¹ Sasha Murrell,^{2,3} Maneerat Koohapitagtam,^{1,4} Michael Golden,² Charlotte Julian,¹ Serge Galzi,¹ Marilyne Uzest,¹ Marguerite Rodier-Goud,⁵ Angélique D'Hont,⁵ Marie-Stéphanie Vernerey,¹ Paul Wilkin,⁶ Michel Peterschmitt,¹ Stephan Winter,⁷ Ben Murrell,^{2,8} Darren P. Martin,² and Philippe Roumagnac¹

¹ CIRAD-INRA-SupAgro, UMR BGPI, Campus International de Montferrier-Baillarguet, 34398 Montpellier Cedex-5, France. Email: filloux@cirad.fr

² Computational Biology Group, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town 4579, South Africa.

Oepartment of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.
Department of Pest Management, Faculty of Natural Resources, Prince of Songkla University, Hat Vai campus, Thailand 90120.

⁴ Department of Pest Management, Faculty of Natural Resources, Prince of Songkla University, Hat Yai campus, Thailand 90120.

CIRAD, UMR AGAP, TA A-108/03, Avenue Agropolis, F-34398 Montpellier Cedex 5, France.
 Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, UK.

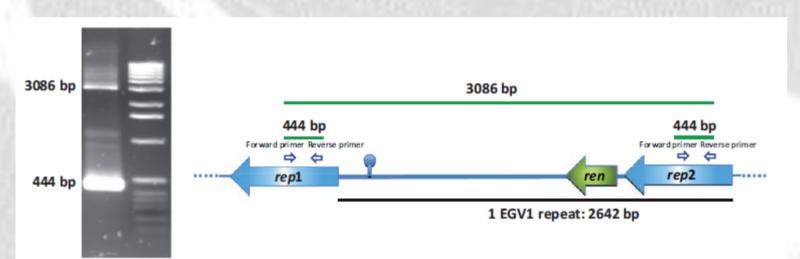
⁷ DSMZ Plant Virus Department, Messeweg 11/12, 38102, Braunschweig, Germany.

⁸ Department of Medicine, University of California, San Diego, La Jolla, CA.

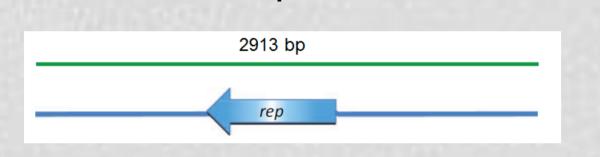
Endogenous viral sequences are essentially 'fossil records' that can sometimes reveal the genomic features of long extinct virus species. Although numerous known instances exist of single-stranded DNA (ssDNA) genomes becoming stably integrated within the genomes of bacteria and animals, there remain very few examples of such integration events in plants. The best studied of these events are those which yielded the geminivirus-related DNA elements (GRD) and the geminivirus-like elements (EGV) found respectively within the nuclear genomes of several *Nicotiana* species (Kenton *et al.* 1995; Bejarano *et al.* 1996; Ashby *et al.* 1997; Murad *et al.* 2004) and various *Dioscorea* spp. of the *Enantiophyllum* clade (Filloux *et al.*, 2015).

Those two new classes of endogenous plant virus sequence are apparently derived from ancient geminiviruses in the genus *Begomovirus*. GRD and EGV sequences likely became integrated millions years ago. Interestingly, we found evidence of natural selection actively favouring the maintenance of EGV-expressed replication-associated protein (Rep) amino acid sequences, which clearly indicates that functional EGV Rep proteins were probably expressed for prolonged periods following endogenization.

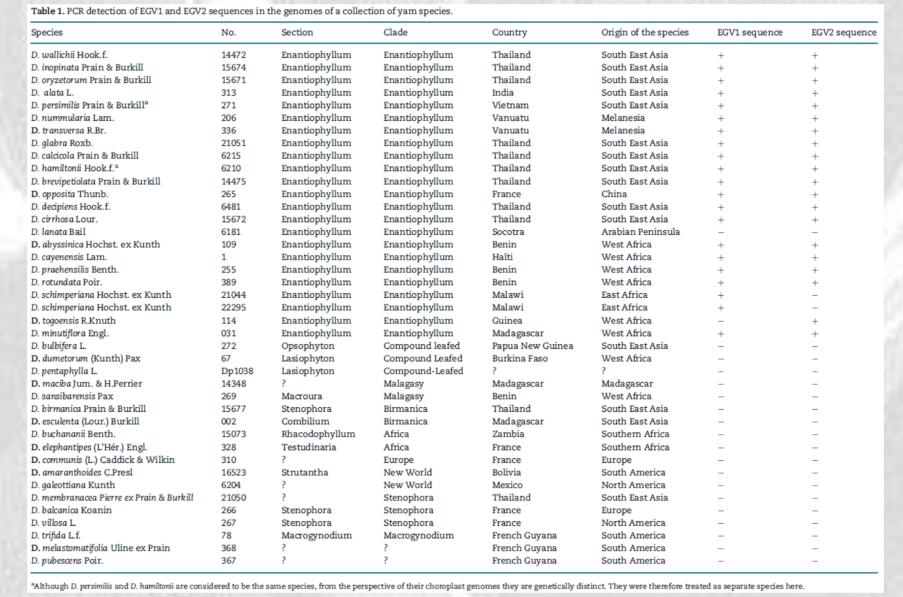
The genomes of many yam (Dioscorea spp.) species contain transcriptionally active endogenous geminiviral
sequences that may be functionally expressed



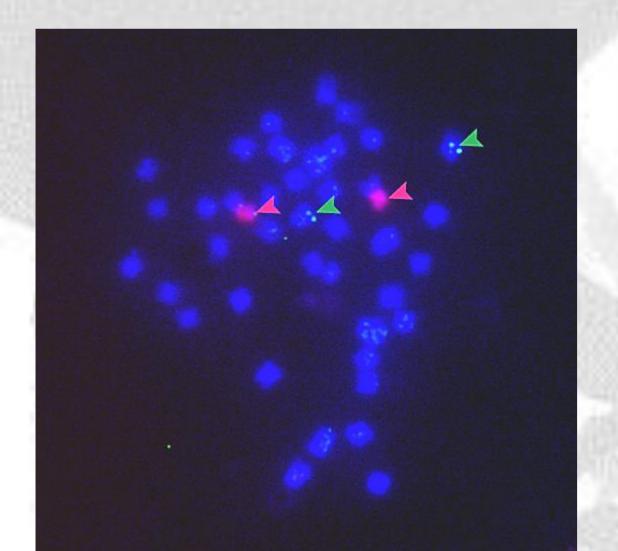
Characteristics of the EGV1 sequence: partial tandem repeat, TAATATTAC, *rep*, *ren* and absence of *cp*



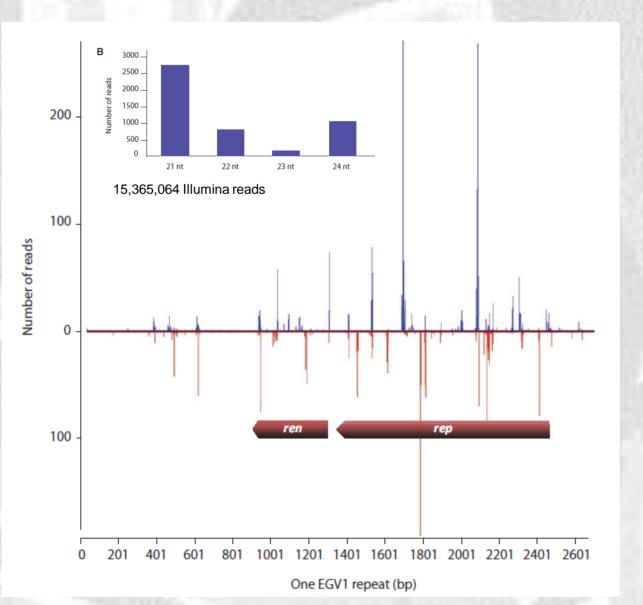
Characteristics of the EGV2 sequence: only *rep* (absence of *cp*, *ren* and *rep*)



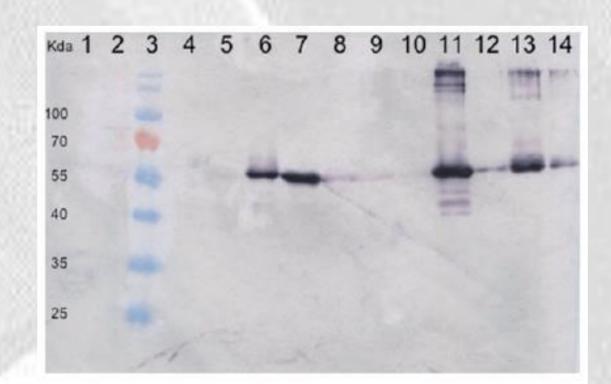
Distribution of EGV1 and EGV2 sequences among members of the *Dioscoreacea* family



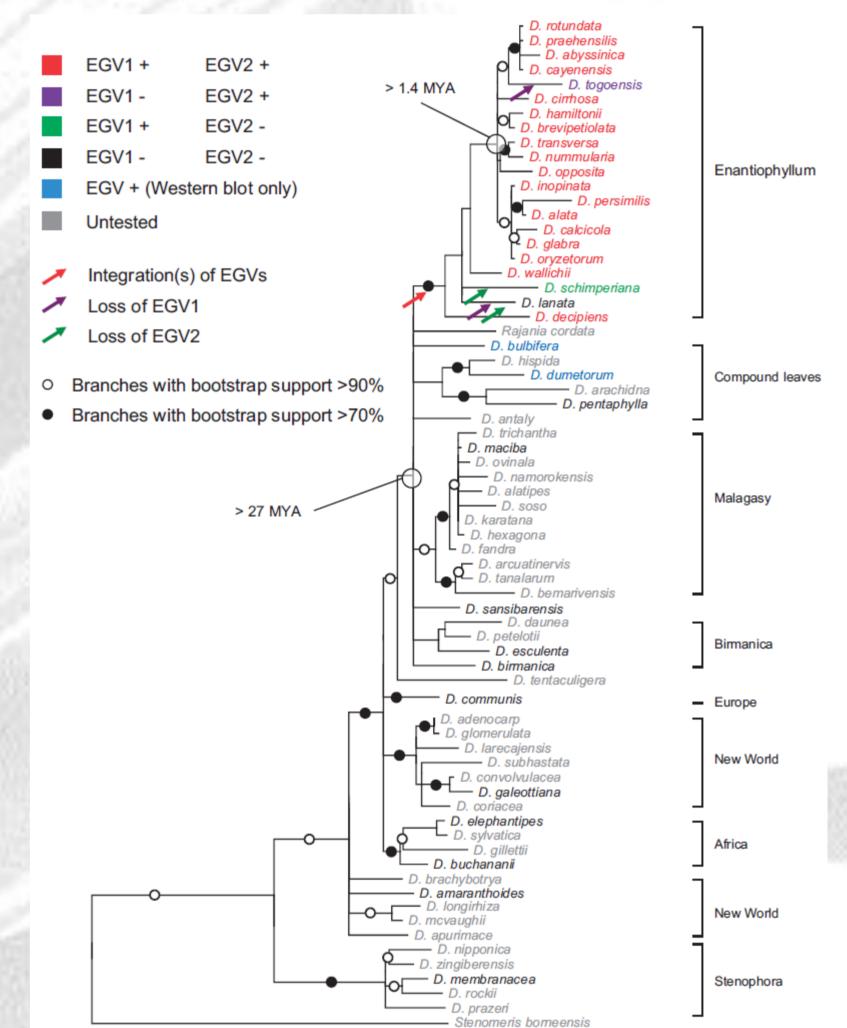
FISH on *D. alata* chromosomes with a 2.6kb EGV1 probe (green arrows) and a rDNA 45S probe (pink arrows)



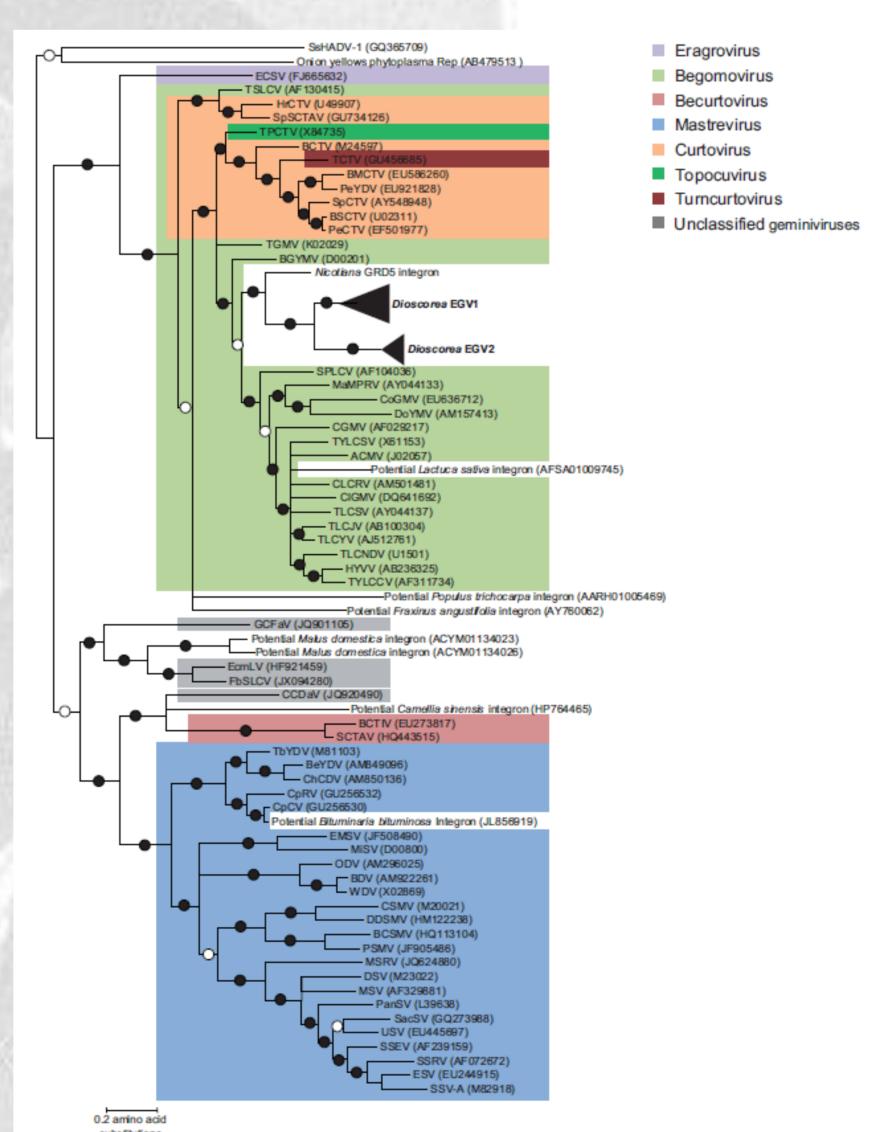
siRNA mapping on EGV1 sequence suggests that *rep* and *ren* genes are transcriptionally active



Western blot on the total protein extracts of several *Dioscorea* species using an antibody directed to Rep peptide of EGV1



Maximum likelihood tree describing the evolutionary relationships between clades of *Dioscoreacea* based on *rbcL* and *matK* nucleotide sequences



Phylogenetic relationships between EGV1, EGV2 and the geminiviruses

New *D. alata* sequence resources available to further explore yam EGVs

- > Draft genome of *D. alata* now available on GenBank (CZHE02000000)
- Four *D. alata* BAC clones containing EGVs are now sequenced (CNRGV, INRA, Toulouse, France)

Several other plant species contains EGVs

We recently found using *in silico* searches that other EGVs are included within complete or draft genomes of various plant species, including apple tree (*Malus domestica*), black cottonwood (*Populus trichocarpa*), several *Coffea* spp, eggplant (*Solanum melongena*), lettuce (*lactuca sativa*), and Tepary bean (*Phaseolus acutifolius*), which suggests that endogenous geminiviruses may be more common in plant genomes than has previously been appreciated.

References

Ashby, M. K., Warry, A., Bejarano, E. R., Khashoggi, A., Burrell, M., 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 molecular biology, 35(3), 313-321.

Bejarano, E. R., Khashoggi, A., Witty, M., Lichtenstein, C. P. (1996). Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. Proceedings of the National Academy of Sciences, 93(2), 759-764.

Filloux, D., Murrell, S., Koohapitagtam, M., Golden, M., Julian, C., Galzi, S., Uzest, M., Rodier-Goud, M., D'Hont, A., Vernerey, M.S., Wilkin, P., Peterschmitt, M., Winter, S., Murrell, B., Martin, D.P., Roumagnac, P., 2015. The genomes of many yam species contain transcriptionally active endogenous geminiviral sequences that may be functionally expressed. Virus Evol. 1, 1–17. Kenton, A., Khashoggi, A., Parokonny, A., Bennett, M. D., Lichtenstein, C. P. (1995). Chromosomal location of endogenous geminivirus-related DNA sequences in Nicotiana tabacum L. Chromosome Research, 3(6), 346-350. Murad, L., Bielawski, J. P., Matyasek, R., Kovarik, A., Nichols, R. A., Leitch, A. R., Lichtenstein, C. P. (2004). The origin and evolution of geminivirus-related DNA sequences in Nicotiana. Heredity, 92(4), 352-358.







