A Next Generation Sequencing approach to elucidate the existence of ten viral species associated with CSSD in West Africa

Emmanuelle Muller, Andrew Wetten, Joel Allainguillaume, Francis Abrokwah, Koffié Kouakou, Henry K. Dzahini-Obiatey

ISCR- Lima, Pérou- November 15, 2017

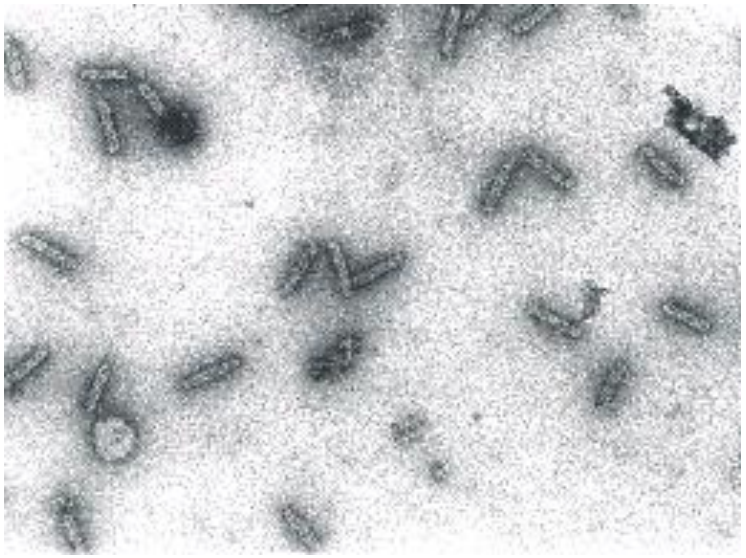
Cacao swollen shoot disease

- Characteristic disease symptoms on *Theobroma cacao* :
  - red vein banding in young leaves
  - shoot, stem and root swelling
  - rounded pods
  - and death of the tree

cirad

Cacao swollen shoot virus (CSSV)

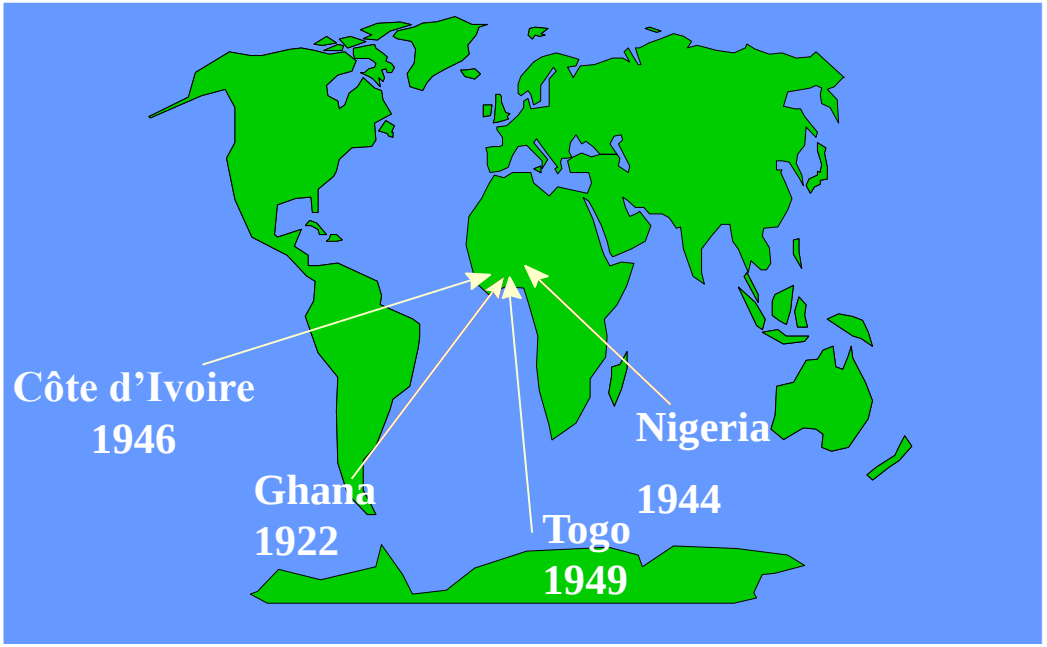
- Family *Caulimoviridae*
  Genus *Badnavirus*
- dsDNA (7-7.3kbp)
- Bacilliform particles (30-150nm)
- Semi-persistent transmission by mealybugs (*Pseudococcidae*)
- No transmission confirmed by seeds



cirad

CSSV epidemics in West Africa: differentiated situations according to the countries

Geographic and historical distribution of the disease

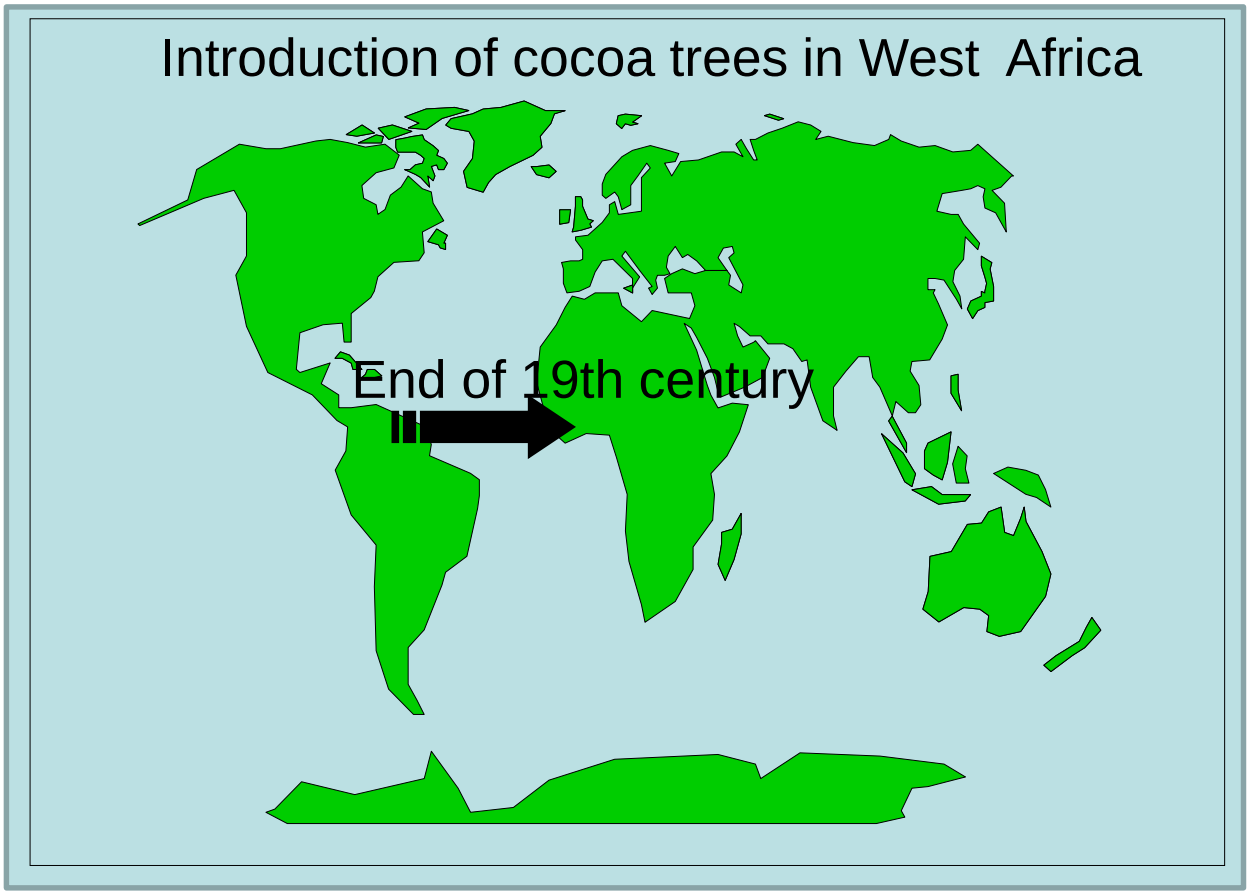


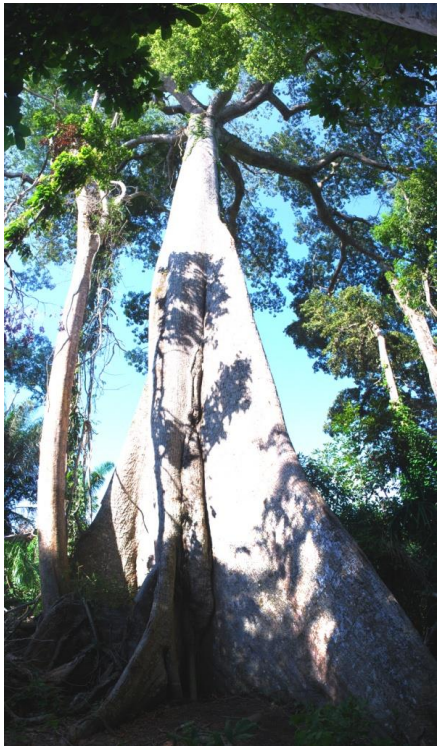
In Ghana: the disease **diffused rather quickly** almost everywhere,
New outbreaks in the West

In Togo: the disease **diffused only** in the south producing area (Kloto) until 2000,
New outbreaks since 2000 in the main producing area (Litimé)

In Côte d'Ivoire: The disease has **not diffused** from outbreaks in the East until 2002
New outbreaks in the West central area and now everywhere

cirad



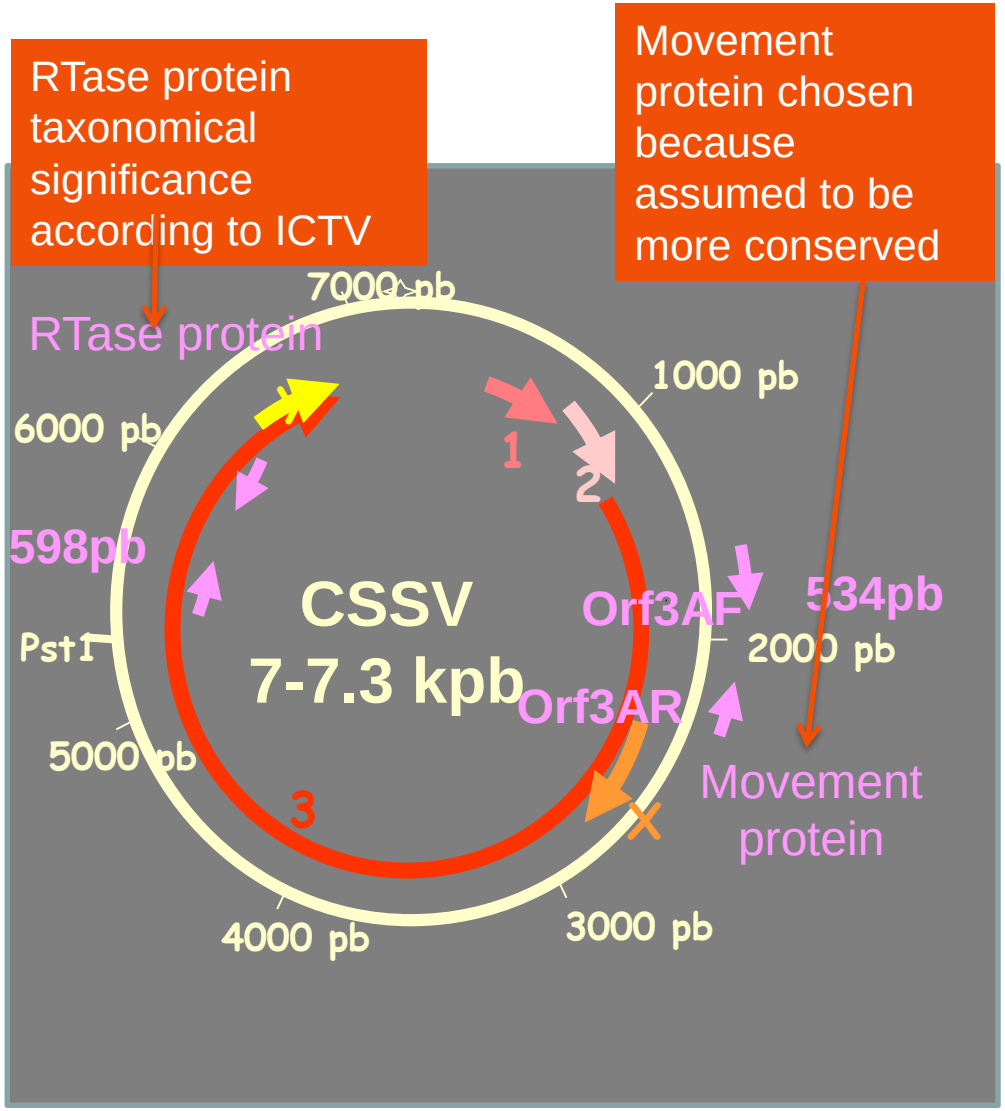


Probable origin : host shift (s) from indigenous reservoir plants
(*Sterculiaceae* : Cola species, *Bombaceae* : *Adansonia digitata* -baobab-, *Ceiba pentandra* -kapoktree-)

cirad

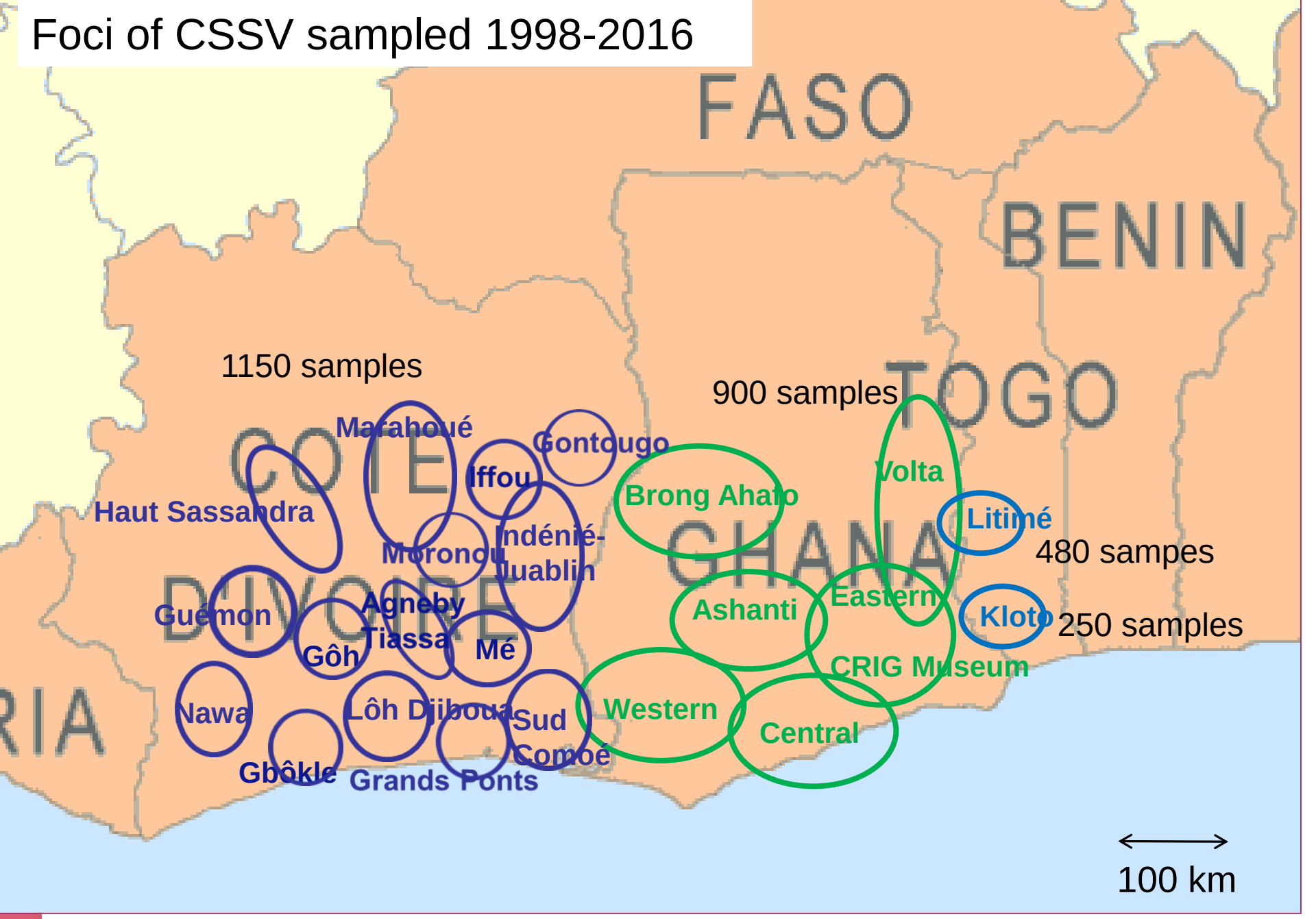
CSSV Variability studies

- Prospections over several successive years in Ivory Coast, Togo and Ghana
- Detection of samples with primers specific to CSSV in two genome regions
- Variability by direct sequencing of amplified products (or cloning if mixed infection)
- Sequence alignment and phylogenetic studies



cirad

Foci of CSSV sampled 1998-2016



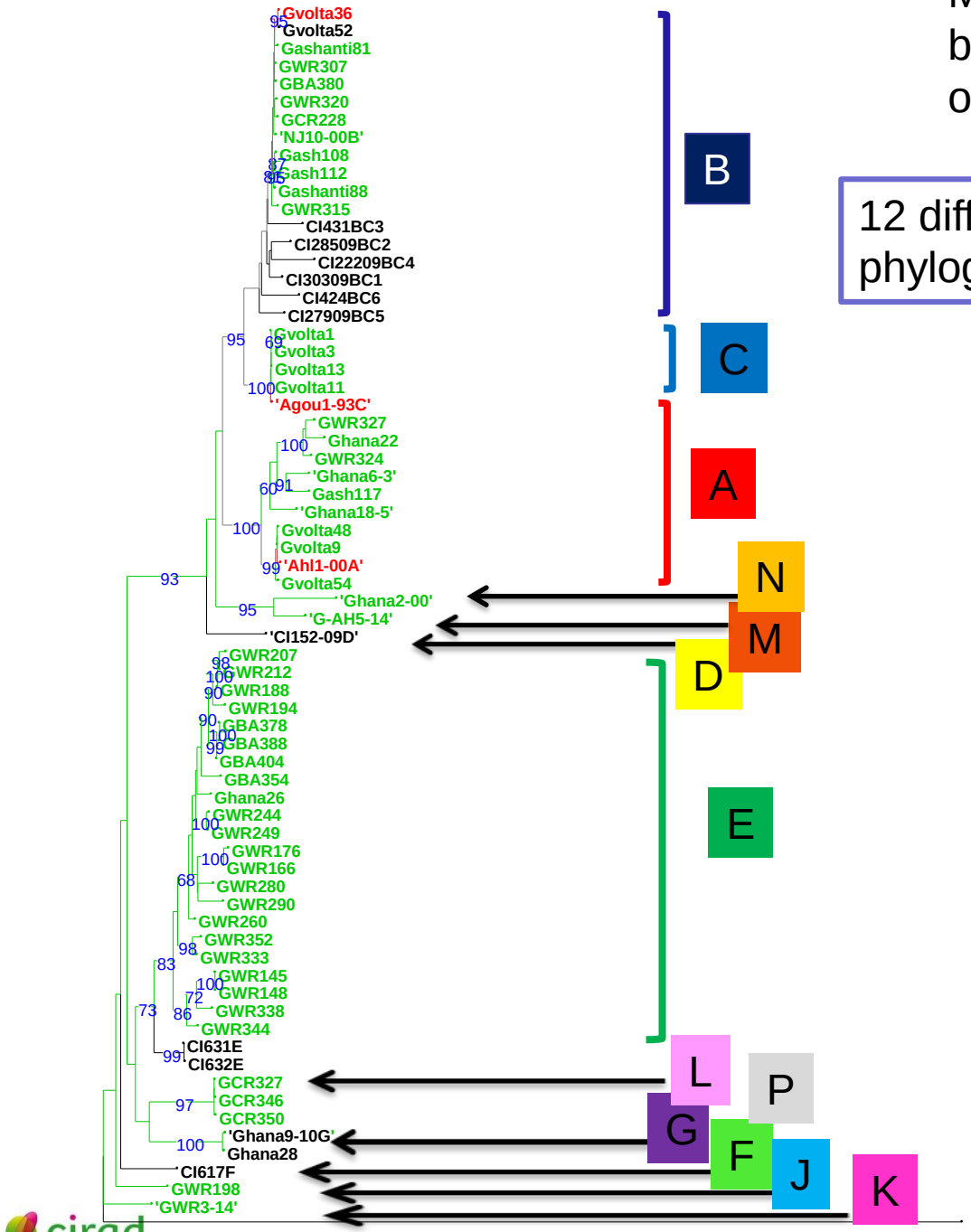
cirad

CRIG Museum diversity (CSSV collection)

- CRIG Museum established in the 1940s when the first surveys began
- 74 isolates analysed

cirad

Maximum likelihood phylogeny, based on alignment of the first part of ORF3 (movement protein)



12 different phylogenetic groups

Origin of isolates:

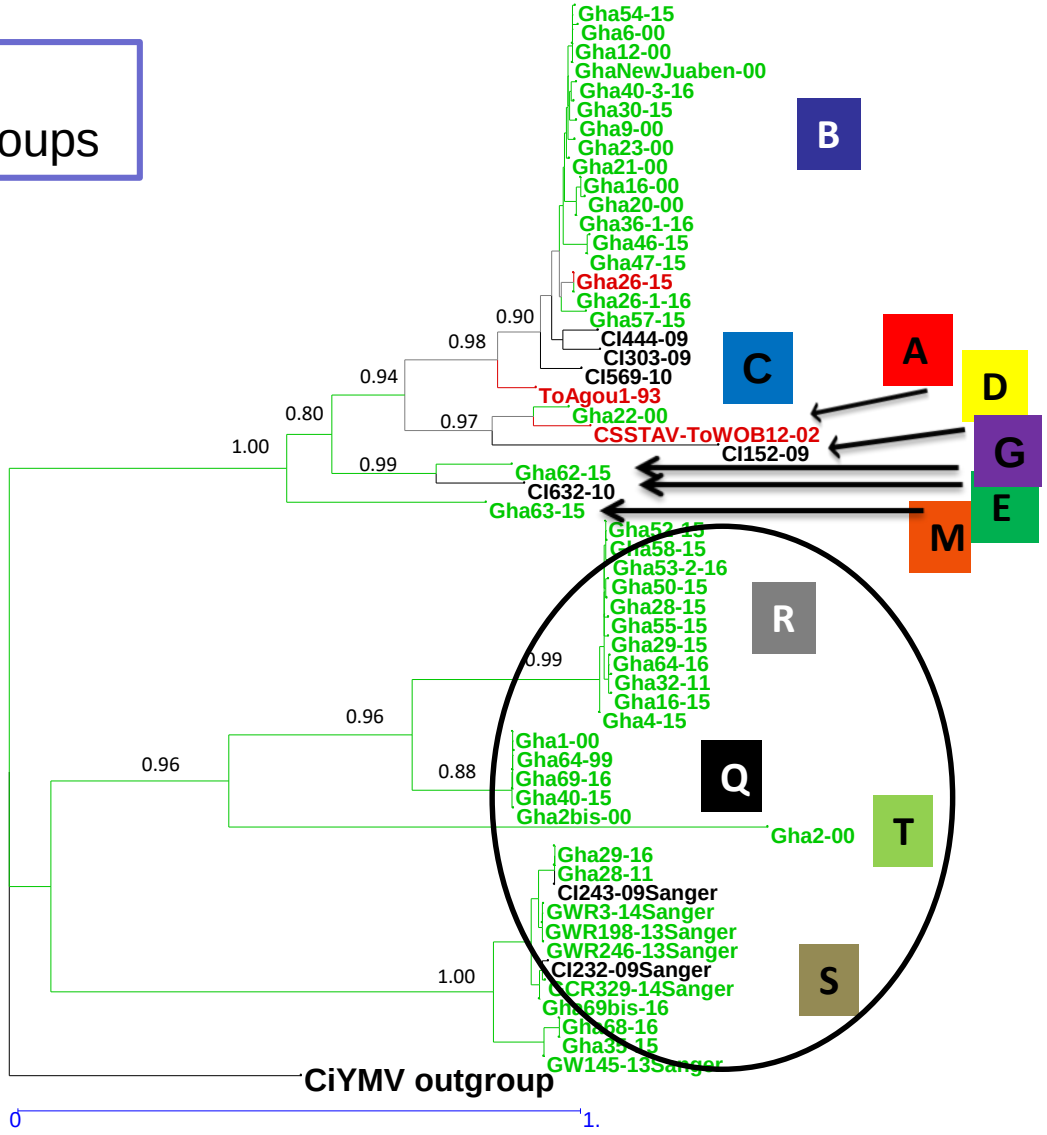
Côte d'Ivoire
Togo
Ghana

cirad

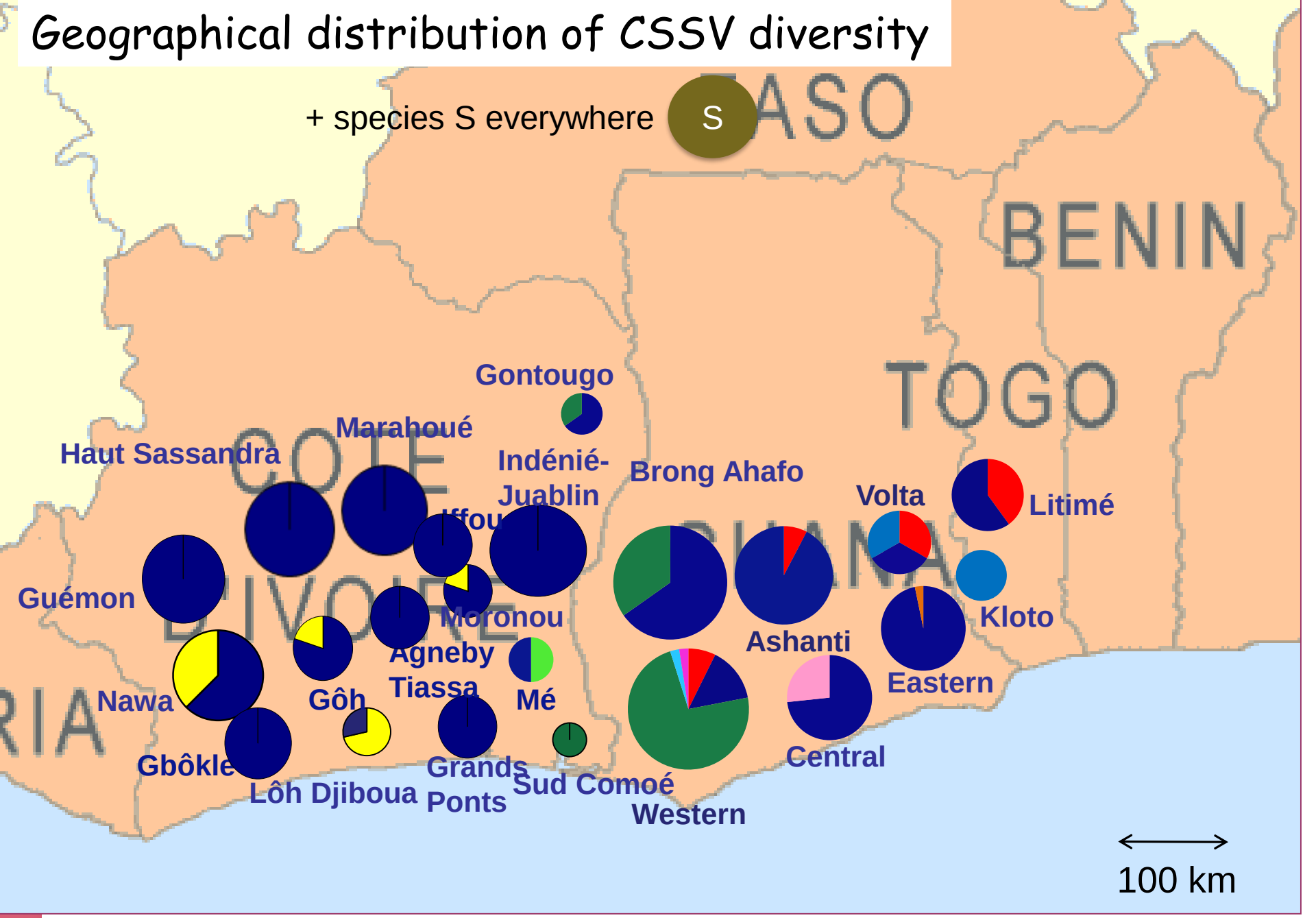
Maximum likelihood phylogeny of CSSV sequences, based on alignment of the RTase region in ORF3

4 additional phylogenetic groups

Côte d'Ivoire
Togo
Ghana



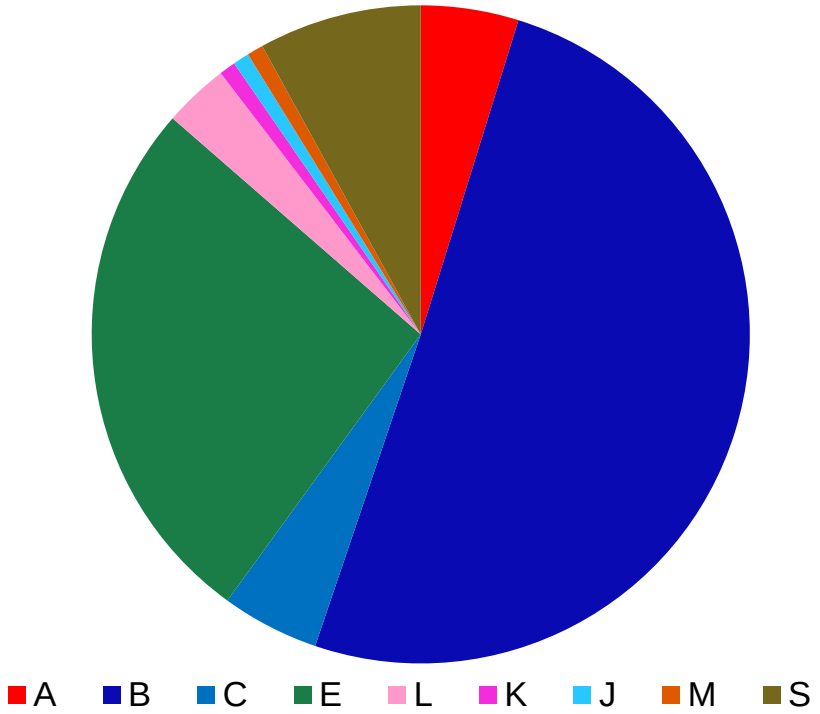
Geographical distribution of CSSV diversity



A B C D F E J K L M

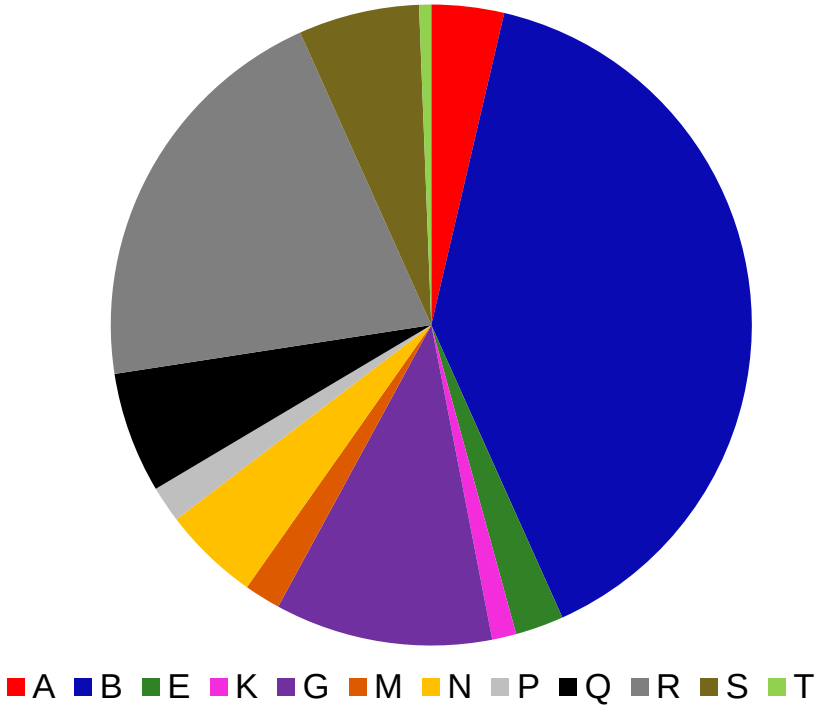
Field CSSV diversity versus CRIG Museum diversity in Ghana

830 samples from the cacao farms of the six cacao growing regions



Cacao farms
9 groups

70 isolates (and 123 samples) analysed from CRIG (with 40 mixed infection)



Tafo CRIG Museum
12 groups

cirad

Conclusions of the CSSV diversity work

- High variability of the CSSV populations in West Africa : a complex of different viral species is responsible for the Cacao swollen shoot disease

It remains to describe viral diversity in Nigeria

- One ubiquitous group, B, the other groups disseminated more locally.

- Some groups are only detected in the CRIG Museum (G, N, P, Q, R)

Some groups only present in the cacao farms (C, L, J absent from the Museum, E present but underrepresented in the Museum) probably corresponds to recently emerged groups

- New outbreaks in the West of Ghana and in the West central area of Côte d'Ivoire associated with new groups, respectively E and D.

Emerging issues on the CSSV diversity described

- A better understanding of the extent of CSSV diversity in terms of different taxonomical species in the different countries

- Improving the detection primers because a lot of symptomatic samples are not detected (depending of the region and of the ages of the leaves)

  - Need to amplify CSSV sequences without a priori in all symptomatic samples.
  - Need to obtain the complete genomes corresponding to the different groups/species.

The next generation sequencing as Illumina sequencing was able to help to obtain more new sequences and CSSV complete genomes

Results obtained from the Illumina sequencing

Illumina HiSeq sequencing was done on DNA obtained from purified cacao leaves infected by CSSV (by Fasteris SA Switzerland),

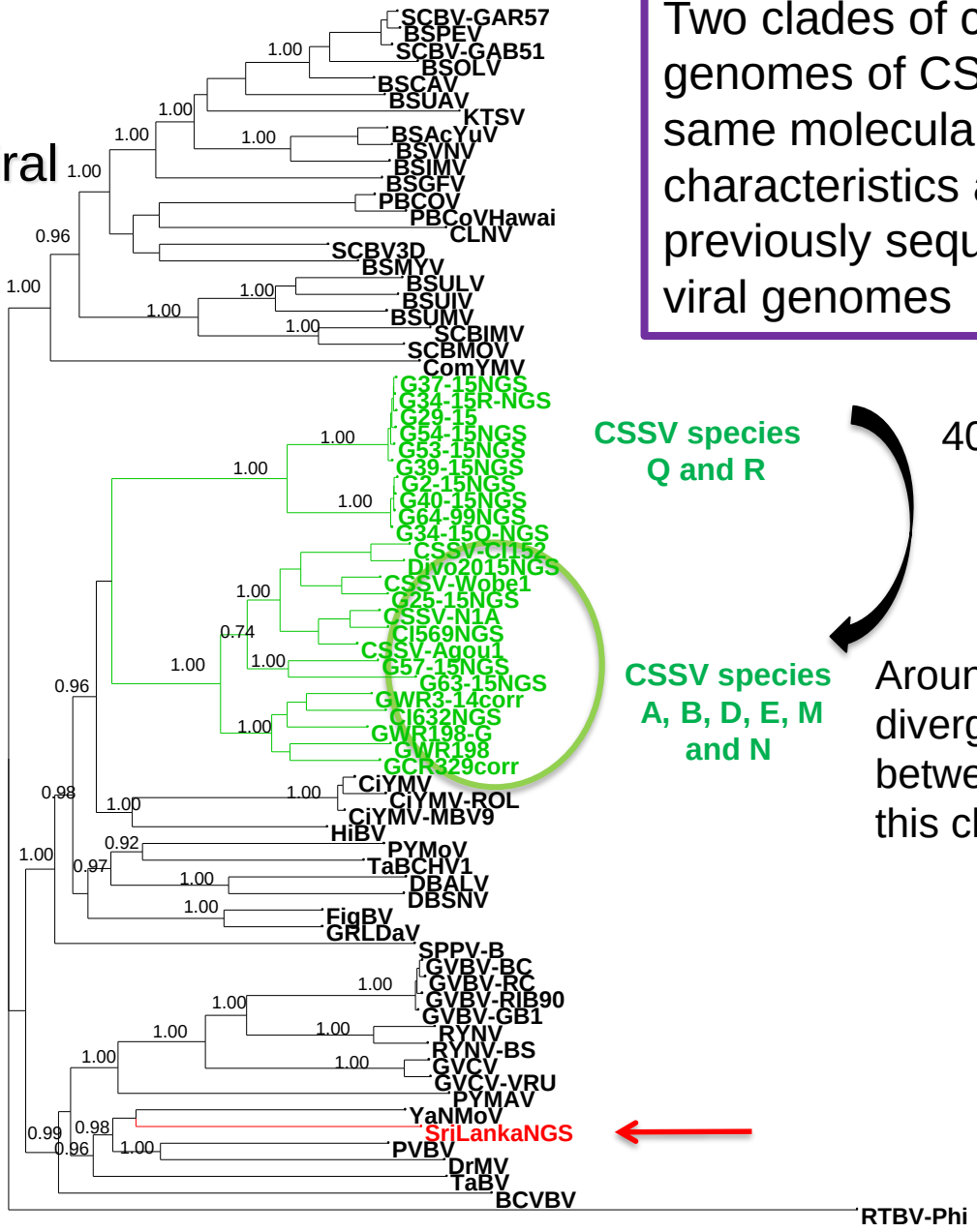
- Samples containing new groups of CSSV (and potentially new species) in order to reconstruct complete genomes
- Samples impossible to detect with diagnostic primers (pooled field samples)

Reads from Theobroma cacao were removed, and the remaining reads are assembled with a pipeline including SPAdes software

21 complete genomes obtained, of which 18 belong to groups containing only partial sequences.
Construction of new phylogenies for RTase protein, and complete genomes

cirad

Maximum likelihood phylogeny of the Badnavirus genus, based on complete viral sequences



Two clades of complete genomes of CSSV with same molecular characteristics as previously sequenced viral genomes

CSSV species Q and R

40 % divergence

CSSV species A, B, D, E, M and N

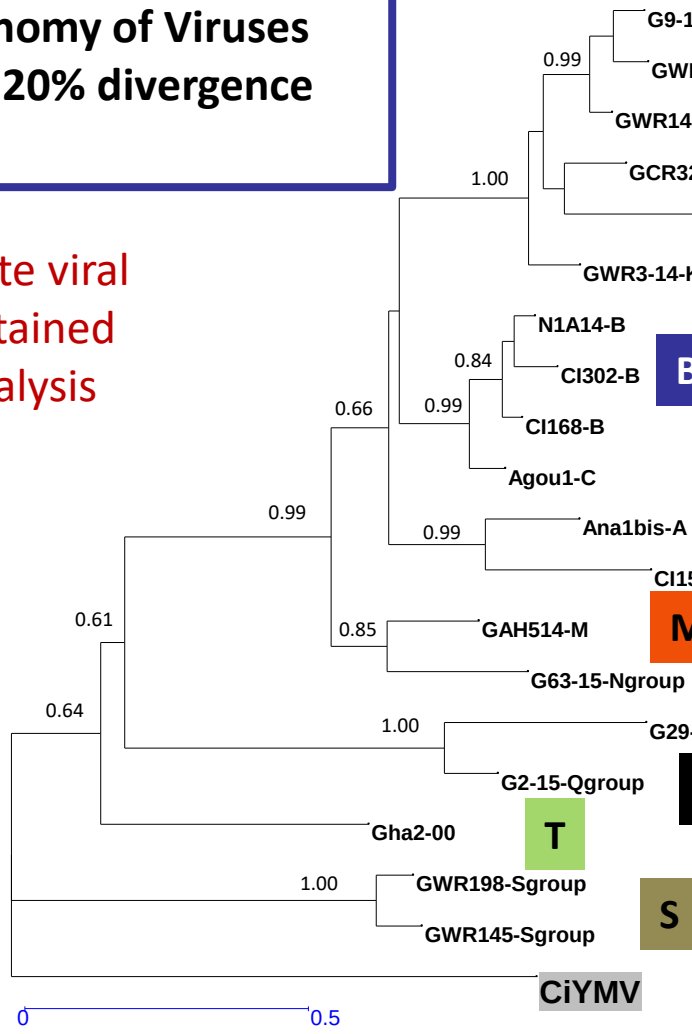
Around 25 to 30% divergence between species of this clade

cirad

Maximum likelihood phylogeny of CSSV sequences, based on alignment of the RTase region in ORF3

10 species according to International Committee of Taxonomy of Viruses recommendations (20% divergence threshold)

New complete viral genomes obtained with NGS analysis



E (subgroups F, K, J, L, G) = CSSCEV ← 5

B (subgroups B, C) = CSSV ← 1

A= CSSTAV ← 1

D = CSSCDV ← 1

M ← 1

N ← 1

R ← 6

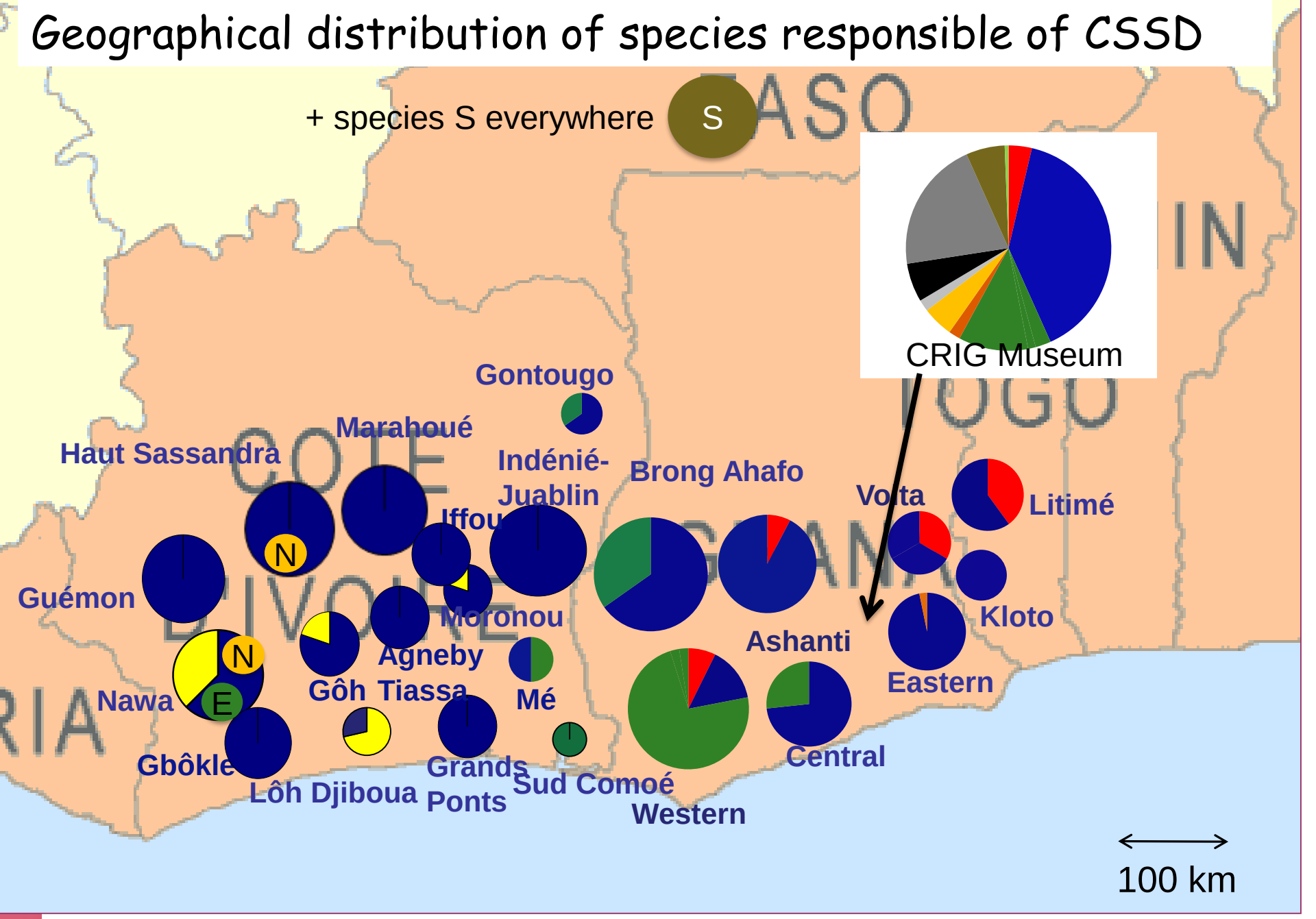
Q ← 4

T

S

Four species T R S Q very different from the others (<60% identity)

Geographical distribution of species responsible of CSSD



Conclusions of NGS work

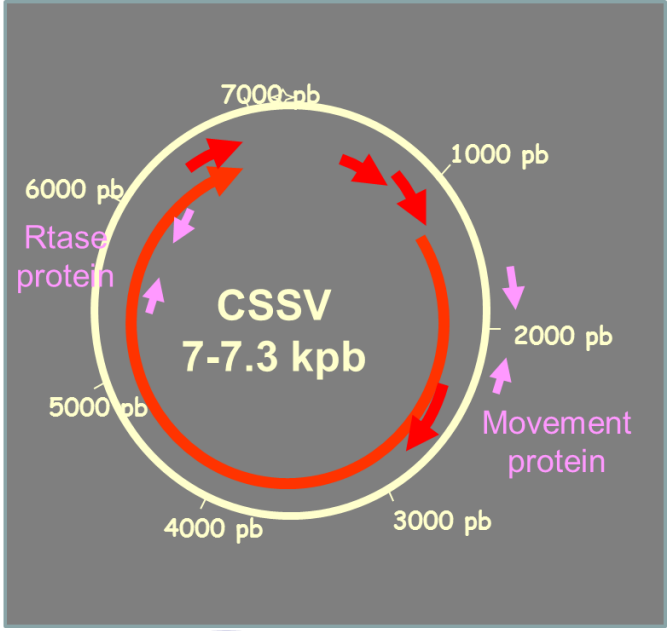
- no new groups identified by the NGS technology but obtention of new complete genomes corresponding to groups already identified with partial sequences

- identification of mixed infection in the CRIG Museum and in the field (partial and complete sequences)

- Identification of viral sequences in some field samples where the PCR detection was not possible (P and N in some Côte d'Ivoire samples)

- The results obtained from NGS technology suggest that the problem of detection is more probably due to the low sensitivity of the PCR protocol

Design of new polyvalent primers for the detection of the 10 viral species in the RTase protein



primers	ORF3A movement	RTase specific	RTase Polyvalent	RTase Species S
samples detection	50%	66%	82%	51%

- More polyvalent primers compared to the previously designed in movement protein
- Cleaning and concentration of DNA to be improved for a better efficiency of the PCR detection (Yield of extraction, RCA step, purification step, more efficient Taq Polymerase)

cirad

Conclusions

➢ The Cacao swollen shoot disease is caused by a **complex of 10 different species** as other diseases caused by badnaviruses (*Banana streak virus*, *Dioscorea bacilliform virus*).

➢ From the high variability described on CSSV populations compared to the very short evolutionary history of CSSV on cocoa trees and the differential distribution of the CSSV groups in the different regions, we suggest **parallel emergences of the disease** in at least two countries (**CSSCDV species in Côte d'Ivoire, E species in Ghana**).

Perspectives

- Understand the role played by the different viral species in epidemics (molecular groups described in the new outbreaks, aggressiveness of the different isolates inside each species, species in coinfection).
- Obtention of full genomes of the two species S and T (specially S, ubiquitous in the field as the first described CSSV species -B-)
- Screening of replanted cocoa material against the viral species present in the corresponding country or area
- Investigate the role of the potential alternative hosts in disease spread and the viral diversity inside these hosts

cirad

Ackowledgements

UMR BGPI Biologie et Génétique des Interactions Plante-Parasite

CNRA CENTRE NATIONAL DE RECHERCHE AGRONOMIQUE

Koffié KOUAKOU
Brice SERIKPA
Ismaël KEBE

cirad LA RECHERCHE AGRONOMIQUE POUR LE DÉVELOPPEMENT

INRA

Montpellier SupAgro

Ghana Cocoa Board
Poised to Maintain Premium Quality Cocoa

CRIG, Ghana

Francis ABROKWAH
Henry K. DZAHINI-OBIATEY

ITRA- CRA/F Kpalimé Togo

Komlan WEGBE
Angèle MISSISSO

CAOBISCO Chocolate, Biscuits & Confectionery of Europe

FCC

eca european cocoa association

World Cocoa Foundation

cirad