



# Task 15 : Targeted Musa Genome Sequencing and Frame Map Construction

**Collaborators:** F. Bakry<sup>1</sup>, F.C. Baurens<sup>1</sup>, J. Dolezel<sup>2</sup>, P. Heslop-Harrison<sup>3</sup>, I. Hippolyte<sup>1</sup>, T. Matsumoto<sup>6</sup>, M. Rouard<sup>4</sup>, N. Roux<sup>4</sup>, M. Souza<sup>5</sup>, T. Sasaki<sup>6</sup>

<sup>1</sup> CIRAD, <sup>2</sup> IEB, <sup>3</sup> University of Leicester, <sup>4</sup> IPGRI, <sup>5</sup> EMBRAPA, <sup>6</sup> NIAS

## The Project



Bananas (*Musa* spp.) are one of the developing world most important food crop and are grown in more than 100 countries throughout the tropics and sub-tropics.

Productivity is declining in many regions. Moreover the production systems that depend on agrochemicals and irrigation are not sustainable.

Cultivars are mostly sterile and parthenocarpic triploids (2n=3x=33), originating from two wild diploid species: *Musa acuminata* (A genome) and *Musa balbisiana* (B genome).

Despite the importance of this culture, basic genetic and genomic tools are missing.

### Objectives:

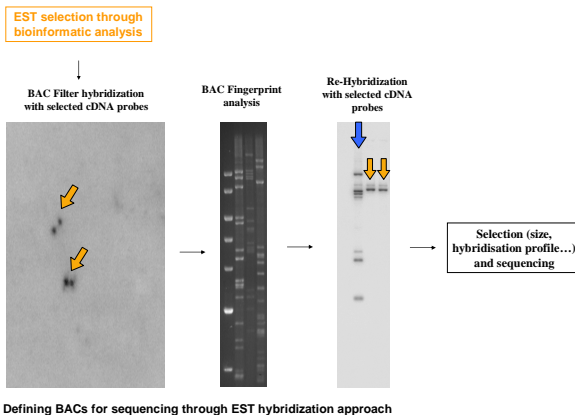
- To obtain a genetic map with 11 linkage groups corresponding to the number of chromosomes of *Musa acuminata*.

The strategy is to produce an F1 population with distant and heterozygote parents. Mapping will be performed with both neutral and derived EST SSR markers.

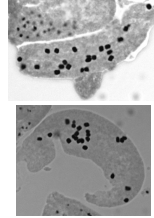
- To develop new genome material.
- To pursue the sequencing (cDNA and BAC libraries).
- To carry out analyses using genetic, genomic and expression data.

## Genomic Resources and Sequencing

- 144x384 well plates of the MA4 BAC library have been pooled.
- 17 BACs bearing genes mainly related to abiotic and biotic stresses have been sequenced. ESTs have been used as probe to identify the BACs to sequence (cf. figure below).
- 2 new cDNA libraries from *Mycosphaerella* infected leaves were produced
- 6132 clones from cDNA libraries obtained from leaves of *Musa acuminata* 'Calcutta 4' infected by *Mycosphaerella fijiensis* were sequenced.
- 614 repetitive DNA sequences, isolated from Calcutta4, were analyzed
- The A and B genomes are being compared, using the highly conserved *adh1* gene and one RGA locus to identify BACs for sequencing. 8 BACs are being sequenced.



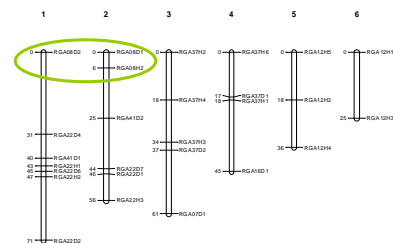
## Mapping population



The BORLI population (*M.a.* 'Borneo' x *M.a.* 'Pisang Lilin') was chosen as mapping population mainly because of the great heterozygosity of both parents (76 % each), the population size (260) and the date of availability of the material.

The root tips of 52 progenies were observed to estimate the chromosome number. 62% shows the expected result (22K) whereas the rest of them presents some additional chromosomes (23K or 24K) probably due to translocations, which are frequent in *Musa*.

The population have been mapped with SSR and RGA (Resistance Gene Analogs) markers. The genetic map is in progress (cf. below)

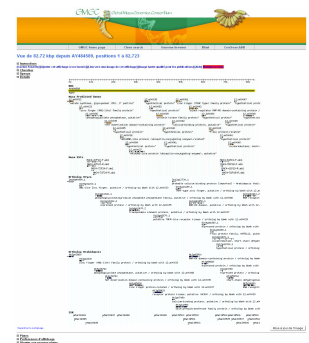


Preliminary map of 12 banana RGAs onto the BORLI population (excluding the 8 most disturbed markers, Kosambi LOD2.0) 12 probes with single or multiple bands used for genotyping 120 individuals mapping population.

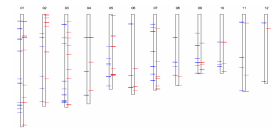
## Bioinformatics

The obtained results are analysed and stored in different databases and tools accessible via the Global Musa Genomics website (<http://www.musagenomics.org>).

**Sequencing status database:** This database intends to keep up to date the list of BACs clones that was, is or will be sequenced. Some important information like the size, date, probes etc. is also available.



**Musa Genome browser:** This tool is used to store and handle genomic information on *Musa*. Functional genes were predicted from the BACs sequenced with the RiceGAAS pipeline. ESTs generated within this project were mapped on the BACs using GMAP and were integrated as well. Other information such as orthologs (rice and *Arabidopsis*) and SSR markers is also present.



**Musa EST mapping onto Rice Genome:** BLASTN (5'-3' EST paired hit ; E value less than 10-20, distances between 5' and 3' EST less than 20kb) identified 192 Musa cDNAs on Rice genome (Build 4 pseudo molecules) (Blue=forward strand hits, Red= reverse strand hits)

Thanks to the support of Generation Challenge Programme, a lot of new resources have been generated on *Musa*. A mapping population is growing up and the 11 linkage groups genetic map is on-going. Genome material is maintained and available at the Musa Genome Resource center. Genomic BAC clones bearing genes of interest have been sequenced and there is now a pretty large EST collection publicly available. Moreover, the Global Musa Genomic Consortium website provides a public access to the outputs of the project.

