

***“Musa genome frame-map
construction and connection with
rice sequence”***

GCP-SP2-project #15



Project team



T. Sasaki (P.I.)
T. Matsumoto



N. Roux
M. Rouard



I. Hippolyte
F.C. Baurens
F. Bakry



J. Dolezel



M. Souza



P. Heslop-Harrison

Musa background information

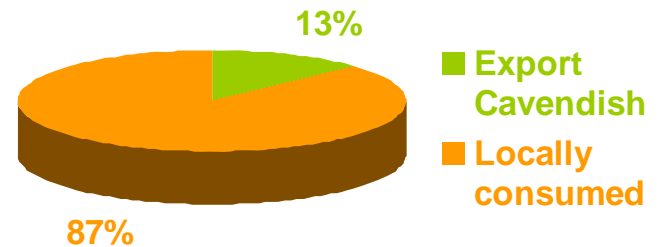


Bananas (*Musa* spp.) are the developing world's fourth most important food crop and are grown in more than 100 countries throughout the tropics and sub-tropics.

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

Most of the cultivars are sterile and parthenocarpic Triploids ($2n=3x=33$), originating from two wild diploid species: *Musa acuminata* (A genome) and *Musa balbisiana* (B genome).

Musa diversity



Edible bananas, or cultivars, are classified according to their genome:

- **AA:** Mostly small sweet bananas
- **AAA-dessert:** eg. Cavendish
- **AAB-dessert:** 'Sweet-acid' flavour
- **AAA-cooking:** EA highland bananas
- **AAB-cooking:** 'true plantains'
- **ABB-cooking:** Starchy flesh

The *Musa* Paradox

Even though *Musa* is...

- The 4th most important basic food crop for 400 Million people
- Reaching 100 millions tons/year over 10 millions hectares
- Grown in 120 countries

Musa is still...

- Under - researched
- Under - funded

Limitation for genomics development in *Musa*

Objectives

1. Consolidate existing publicly available *Musa* genomic resources and make new resources available
2. Establishment of international mapping populations with the view of precision genetic mapping
3. Anchor maps with SSR markers with reference to gene and genome sequences from rice and other models
4. Selective sequencing of BACs bearing genes potentially involved in important traits
5. Establishment of public databases; distribution of clones; training and dissemination of information for use

Musa genomic resources

Genotypes

- *Musa acuminata* 'Calcutta 4' (AA)
- *Musa balbisiana* 'Pisang Klutuk Wulung'
- *Musa acuminata* 'Tuu Gia' (AA)
- *Musa acuminata* 'Grande Naine' (AAA)

BAC libraries

- 17280 clones - 3x genome coverage (C4BAM)
- 55296 clones - 9x genome coverage (MA4)
- 36864 clones - 9x genome coverage (MBP)
- 30700 clones - 5.1x genome coverage
- 55296 clones - 17x genome coverage

cDNA library

- Cloneset (208 x 384-well plates)

- Repetitive DNA clones ('Radka' series)
- Cytogenetic probes (5S rDNA, 45S rDNA, telomere)

Musa Genome Resources Centre (MGRC)

The mission:

- Acquire
- Produce
- Store
- Distribute

Availability (MTA required) :

- Complete libraries
- Filter sets
- Individual 384-microtitre plates
- Single clones

www.musagenomics.org

Musa Genome Resources Centre (MGRC) - Microsoft Internet Explorer

Soubor Úpravy Zobrazit Oblíbené Nástroje Nápověda

Zpět Hledat Oblíbené Média

Adresa <http://www.musagenomics.org/mgrc/index.php>

Musa Genome Resources Centre (MGRC)

Home Overview Resources Ordering site map

Welcome to the *Musa* Genome Resources Centre (MGRC)

The *Musa* Genome Resources Centre (MGRC) was established at the Laboratory of Molecular Cytogenetics and Cytometry of the **Institute of Experimental Botany (IEB)**, Olomouc, Czech Republic in 2003. The Centre is the distribution arm of the **Genomics Consortium** and operates on a cost-recovery basis. The primary aim of the Centre is to support research activities of the *Musa* Genomics Consortium by distributing *Musa* genome resources to Consortium Members and by developing new resources. However, the Centre distributes *Musa* genome resources also to individuals and research organizations outside the Consortium.

This site contains information on currently available **BAC genomic libraries, EST libraries, short insert DNA libraries, repetitive DNA clones, and molecular cytogenetic markers**. Current research and experimental procedures used in the Laboratory of Molecular Cytogenetics and Cytometry of the IEB are described.

Genome resources are supplied to users only under the terms and conditions stipulated by **Material Transfer Agreement (MTA)**, which ensures that the material and information related to it stays in the public domain. For ordering information please read our **ordering information page**.

Please contact the MGRC director, **Dr. Jaroslav Dolezel**. Tel: (+420) 585 205 852). fax: (+420) 585 205 853) with questions about the resources, proposals for scientific collaborations, suggestions for future genomic resources and requests for library copies and commercial licenses.

Also you can contribute. Please send us material (library, DNA clones...etc.) you would like others to use. We will make it available to the scientific community at cost recovery bases. Intellectual property rights will stay with the contributors who will be dully acknowledged and cited as reference.

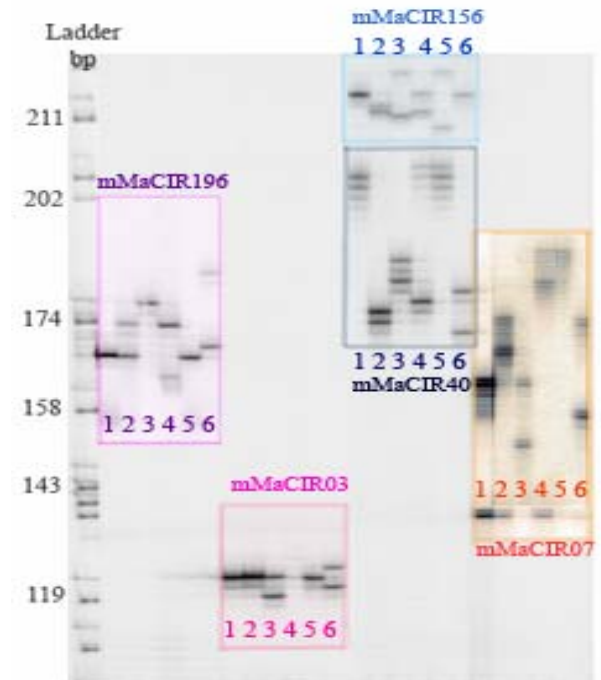
Development and choice of a mapping population

A *Musa acuminata* Pahang doubled haploid (♀) x *M. a.* Pisang Madu (♂)

B *M.a.* Borneo (♀) x *M.a.* Pisang Lilin (♂)

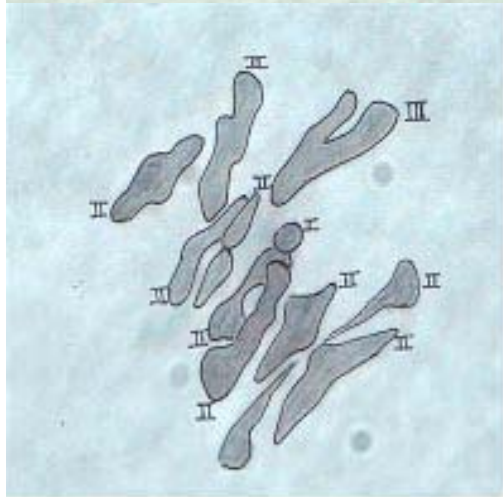
C *M.a.* Malaccensis (♀) x *M.a.* Gabah Gabah (♂)

		Allèle nb	Heterozygosity SSR	Heterozygosity 42 RFLP
1	Pahang HD	21/21 SSR	0%	-
2	P. Madu	34/21 SSR	62%	67%
3	Borneo	86/49 SSR	76%	24%
4	P. Lilin	37/21 SSR	76%	56%
5	Malaccensis	28/20 SSR	40%	33%
6	Gabah Gabah	40/21 SSR	90%	61%



Development and choice of a mapping population

Picture of a Metaphase of Pisang Madu (on one plane)



Interpretative diagram (amplification x 2000)

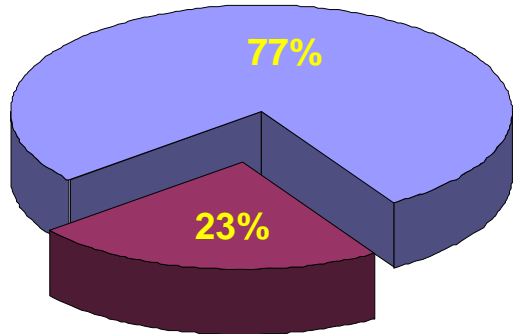
	Translocation breakpoint	Genetic distance 20 SSR loci, simple matching distance	Availability of population
Pahang HD	-	0,775	190 ind. March 2006
P. Madu	1		
Borneo	-	0,78125	280 ind. Dec. 2005
P. Lilin	1		
Malaccensis	-	0,9	200 ind. Dec. 2005
GabahGabah	(P. jari buya) 1		

The important criteria are:

- high heterozygosity,
- high genetic distance between parents,
- numerous individuals,
- the date of availability of the progeny

Characterization of $C_0t-0.05$ library

Total number of clones: 2,688

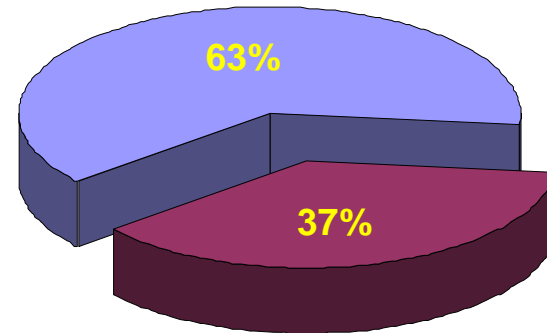


- *Radka* clones ignored
- 614 clones sequenced



- 57 tandem organized repeats
- 21 microsatellite sequences
- 26 palindromic sequences

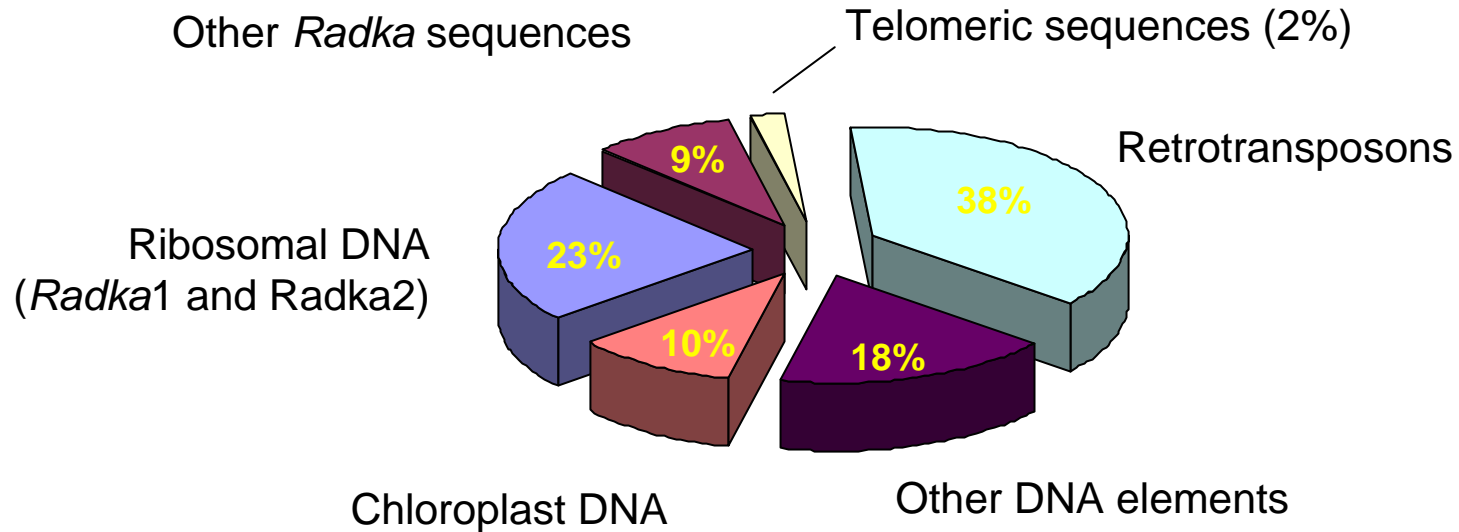
Homology search (GenBank)



- 385 clones not homologous
- 229 clones homologous

Homology to GenBank sequences

229 clones from the $C_0t-0.05$ library



DATA Musa

DATA Musa

Embrapa Genetic Resources and Biotechnology, the Catholic University of Brasilia (UCB) and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) initiated in 2002 a research project entitled "Analysis of the primary structure of the A genome in *Musa acuminata*", financed by the CNPq, with the objective of developing the foundations of a program in genomics and biotechnology for banana in Brazil.

This project resulted in the creation of DATA Musa, today the world's second largest genomic database for banana, composed of structural genomics information (BACs), expressed sequences (ESTs) and Resistance Gene Analogs (RGAs).

Structural genomic sequences have enabled the identification of hundreds of genes, together with selection of promoters of gene expression, which are in the initial phases of validation. Analysis of EST sequences has resulted in identification of 5317 genes, leading to the discovery of several genes of interest for genetic breeding in banana, including those coding for resistance to fungi, bacteria, nematodes and insect pathogens, as well as drought tolerance.

The utilization of information in DATA Musa opens a wide range of possibilities for genetic breeding and directed transgenic applications in the banana crop.

To access DataMusa: <http://genoma.embrapa.br/musa>

Contact:

Dr. Manoel Tebêira Souza Junior - msouza@cnpq.gen.embrapa.br

Embrapa

Empresa Brasileira de Pesquisa Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento
Cruz das Almas, Espírito Santo, CEP 41301-900
Fone: 311 4 404 170; 48 34 760 Fax: 3 40 5172
E-mail: CP

sac@cenargen.embrapa.br

www.cenargen.embrapa.br



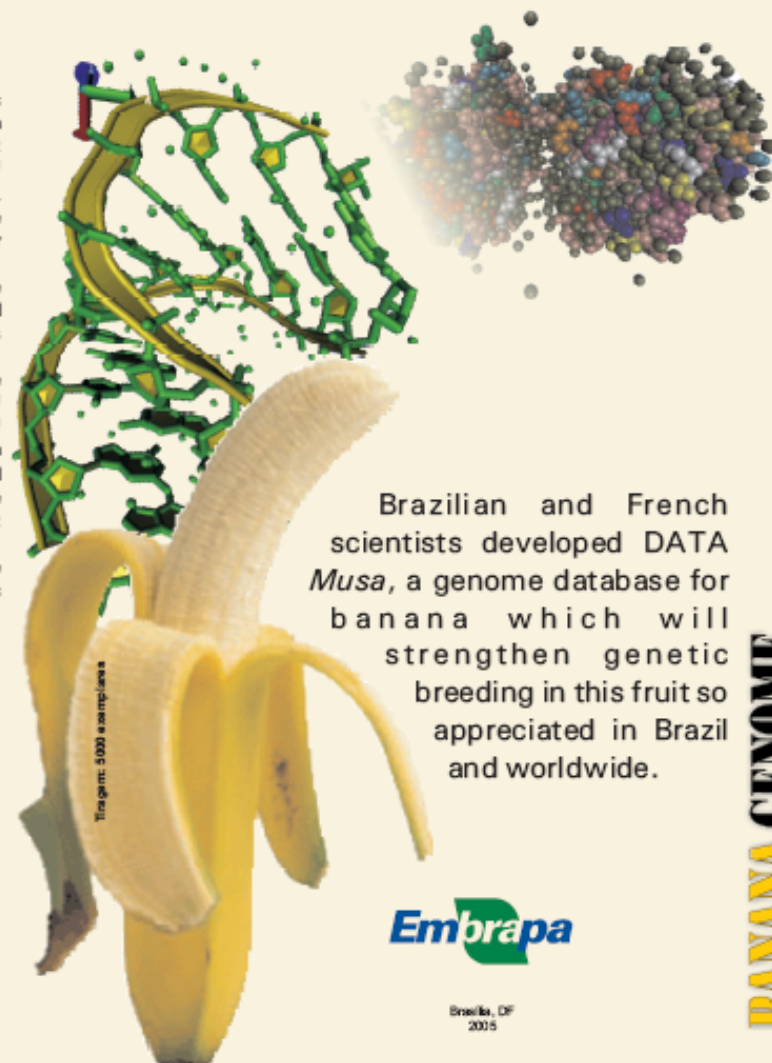
Embrapa
Um Brasil por aí gente

Ministério da Agricultura,
Pecuária e Abastecimento



Embrapa

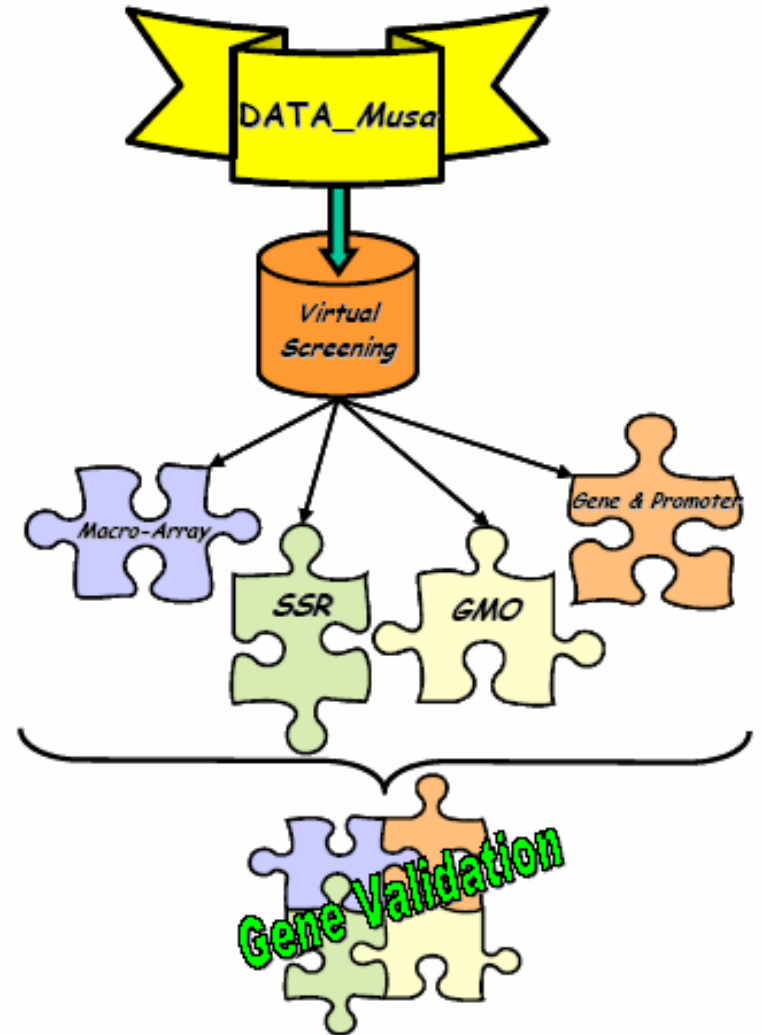
Brasília, DF
2005

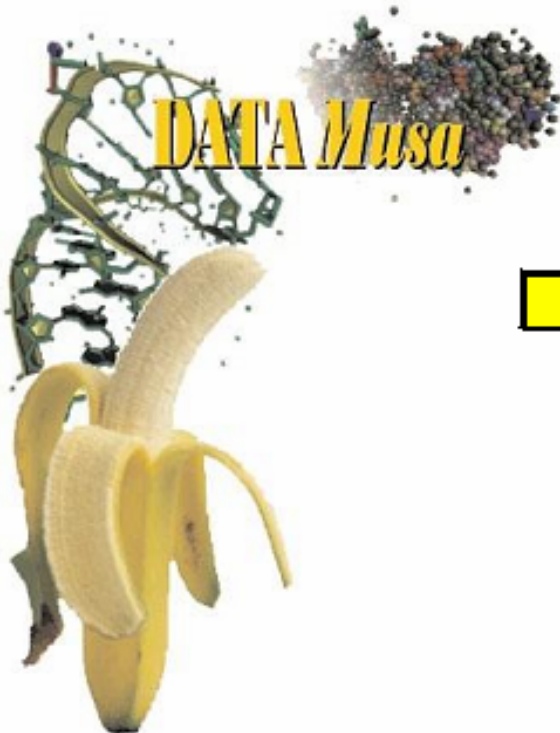


Brazilian and French scientists developed DATA Musa, a genome database for banana which will strengthen genetic breeding in this fruit so appreciated in Brazil and worldwide.

BANANA GENOME

Expressed Sequence Tags





LIST OF cDNA CLONES FROM THE DATA Musa DATABASE FOR BAC CLONES SELECTION
GENERATION CP - Musa MARK

A) Abiotic Stress:

TRAIT	DESCRIPTION	Evalue	CLONE
Drought	Drought Responsive Family Protein	8e-38	MA4LIMFES014A_G03
	Chloroplast Drought Induced Stress Protein	1e-81	MA4LIMFES011B_G02
	Drought Induced S-Like Ribonuclease	8e-57	MUC4RO1004_B09
Salt	Salt Tolerance Protein 5	1e-57	MA4LIMFES012B_G04
Cold	Salt Tolerance Protein (Member of the Constant Zinc Finger Family)	9e-53	MACVLMFELS011D_A07
	Low Temperature Induced Protein	2e-21	MA4LIMFES003C_G06
Heat	Low Temperature and Salt Responsive Protein	4e-15	MA4LIMFES014D_H03
	Class-1 LMW Heat Shock Protein	2e-58	MUC4LH1002_F07
	70 kDa Heat Shock Protein	1e-72	MA4LIMFES014A_E03

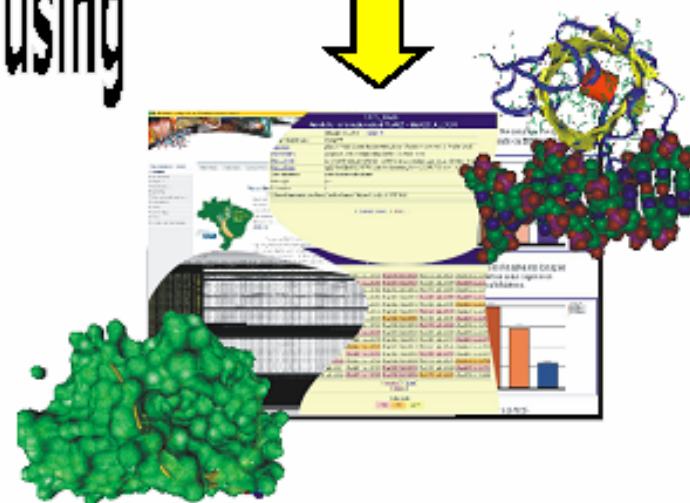
B) Biotic Stress:

DESCRIPTION	Evalue	CLONE
Pathogenesis Related Protein	1e-57	MUC4PE1010_H01
Class III Acidic Chitinase	3e-95	MUC4RO1004_B01
PR1 Pathogenesis Related Protein	1e-63	MUC4RO1021_A10
Catalase	3e-86	MA4LIMFES012A_A03
Ascorbate Peroxidase	3e-98	MA4LIMFES011C_B02
Fe Superoxide Dismutase	6e-47	MUC4PE1014_F07
Beta-1,3-Glucanase	7e-65	MA4LIMFES002A_F07
Hypersensitive-Induced Response Protein	2e-80	MA4LIMFES003B_E09
CES disease resistance protein-like / Leucine-Rich Repeat Family Protein	3e-56	MUC4FL1012_D03
Glutathione Peroxidase	1e-70	MA4LIMFES005A_H05

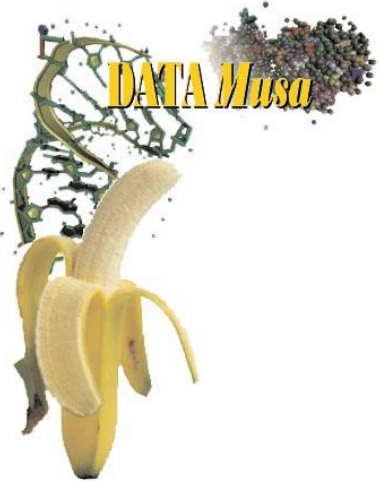
C) Transcription Factor:

DESCRIPTION	Evalue	CLONE
WRKY Transcription Factor 72	1e-46	MA4LIMFES002D_F11
Ethylene-Responsive Transcriptional Coactivator	2e-52	MUC4LH1014_E10

Selection and characterization of BAC clones using banana ESTs from genes related to important traits.



cDNA libraries



- Genome B (*M. balbisiana*) 'Pisang Klutuk Wulung'
- Search for BAC clones harboring genes of interest-
need a 2D or 3D pool of the BAC library

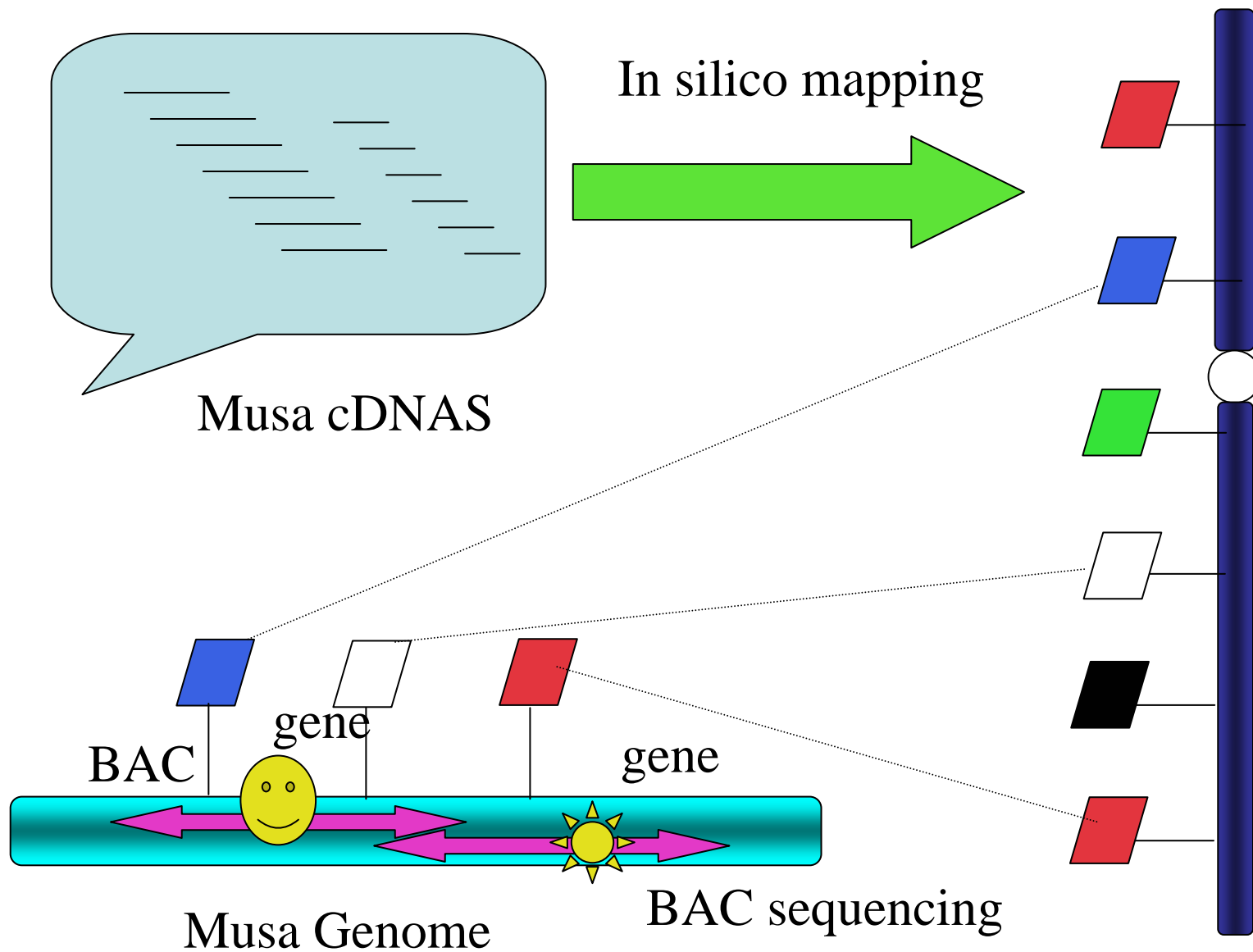
II. Specific Information

1. Abstract of the research activities and expected results

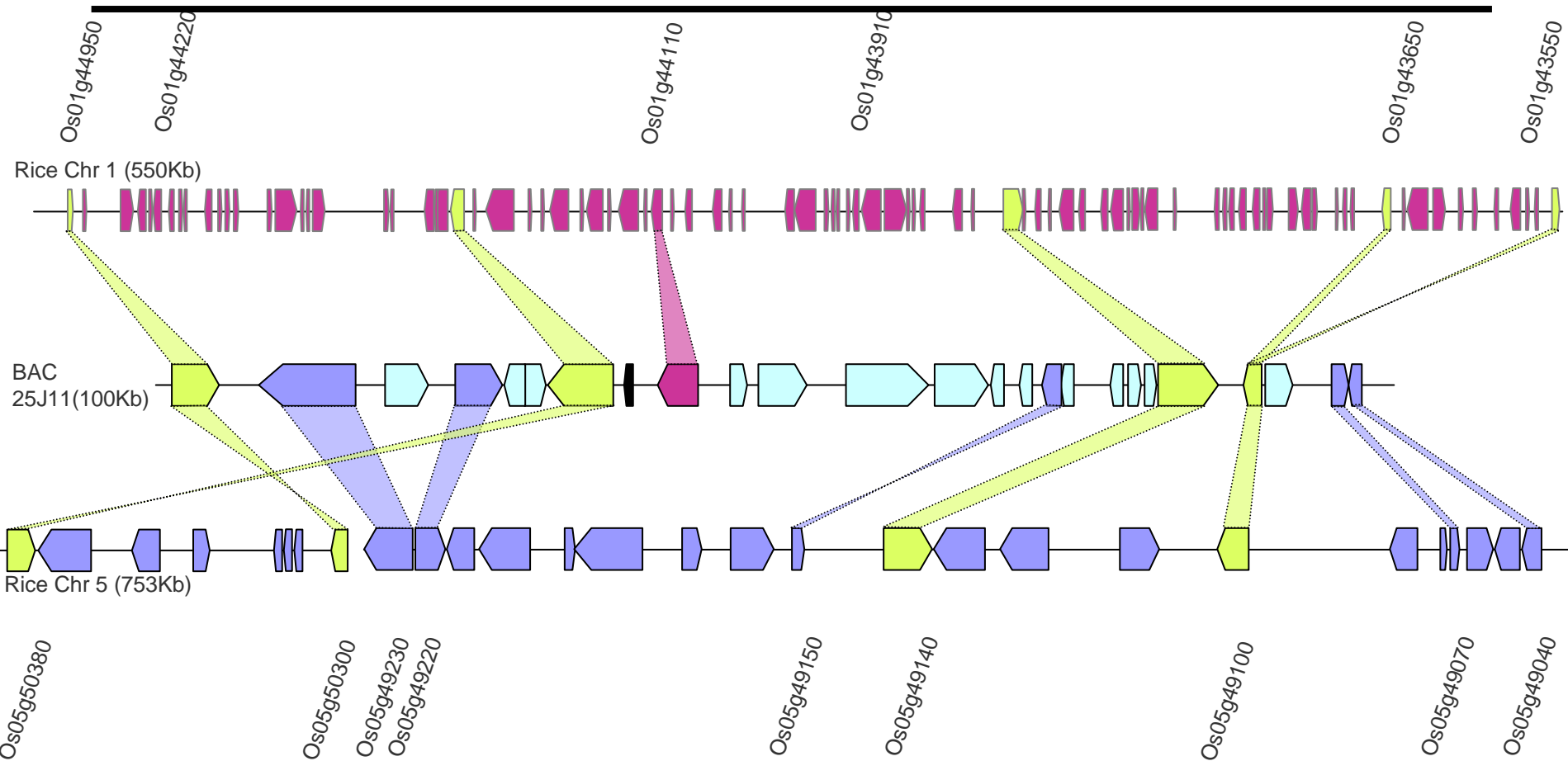
The major goal of this present project is to construct and characterize cDNA libraries from *Musa balbisiana*. Fresh and chronic drought-stressed leaves and roots from *Musa balbisiana* plants (Pisang Klutuk Wulung) will be used to extract total RNA for the production of four cDNA libraries (fresh leaves, drought-stressed leaves, fresh roots, drought-stressed roots). A total of 9,600 ESTs reads (5'-end sequence) being 2,400 ESTs reads per cDNA library will be produced. Bioinformatic analysis (from selection of high quality sequence; assembly of clusters; and BlastX analysis against four public databases – NCBI nr, MIPS, KOG and SwissProt) will be done at the Bioinformatics Lab – Embrapa Genetic Resources & Biotechnology. The chromatograms and the results of bioinformatic analysis will be made available to INIBAP, and a copy will be deposit at DATA *Musa*.

NIAS

Rice Chromosome



Syntenic region between *Musa* and Rice



BLASTP (Altschul et al.1997)
i-ADHoRe (Simillion et al. 2004)



"Musa-Oryza"
GCP-SP2-15 project

Coordination meeting

IPGRI-INIBAP, Montpellier – 6 & 7 July 2005