

# Musa genomic resources

GCP-SP2-project 2005-15

Sasaki T., Matsumoto T., Hippolyte I., Bakry B.,  
Baurens F.C., Sidibe-Bocs S., Dolezel J.,  
Heslop-Harrison J., Souza M., Miller R.,  
Rouard M. and Roux N.



## Objectives:

- ▶ Consolidate existing publicly available Musa genomic resources and make new resources available
- ▶ Establishment of international mapping populations with a view to precision genetic mapping
- ▶ Anchor maps with SSR markers with reference to gene and genome sequences from rice and other models
- ▶ Selective sequencing of BACs bearing genes potentially involved in important traits
- Establishment of public databases; distribution of clones; training and dissemination of information

## Borli genetic map

built from an F1 progeny of *Musa acuminata* 'Borneo' x *Musa acuminata* 'Pisang lilin'

The map is based on :

- 180 SSR genotyped on 180 individuals
- 380 DArT markers genotyped on 92 individuals

The markers are distributed on

- 11 main linkage groups for Borneo and
- 9 linkage groups for Pisang lilin.

Join map4 software was used at LOD 6 with Kosambi function.



# Publicly available Musa genomic resource

## cDNA libraries

*Musa acuminata* (AA) 'Calcutta 4'  
two libraries from *Mycosphaerella fijiensis* infected leaves  
(seq. total 6,132 ESTs)

*Musa balbisiana* (BB) 'Pisang Klutuk Wulung'  
two libraries from leaves and roots  
( seq, total 4,800 ESTs)  
two new libraries under osmotic stress will be constructed

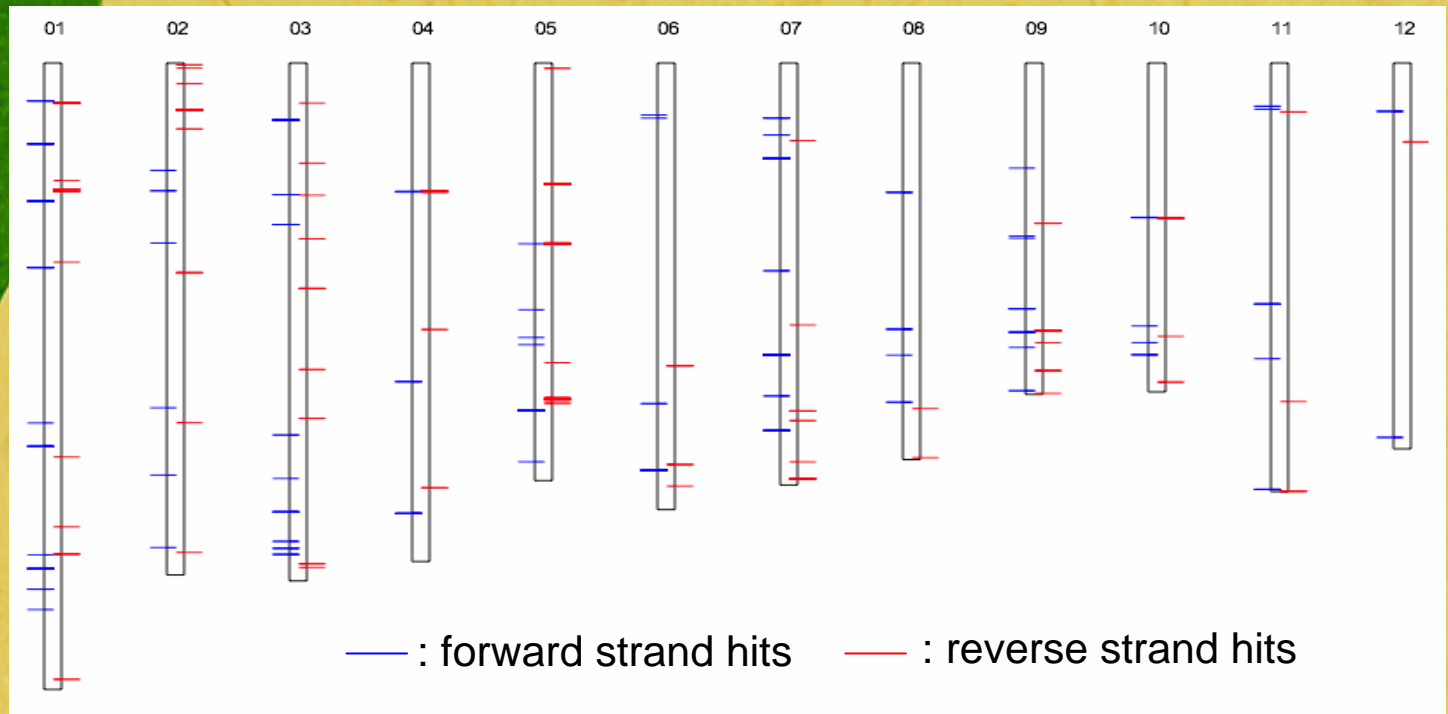
## BAC libraries

17280 clones - 3x genome coverage (C4Bam, calcuta4)  
55296 clones - 9x genome coverage (Ma4,calcuta4) **3D Pool**  
36864 clones - 9x genome coverage (MbP, PKW) **3D Pool**  
30700 clones - 5.1x genome coverage(TG BIBAC, Tuu Gia)  
55296 clones - 4.4x genome coverage(MaC, Grande Naine)  
DH-Pahang(AA) (under const.)

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



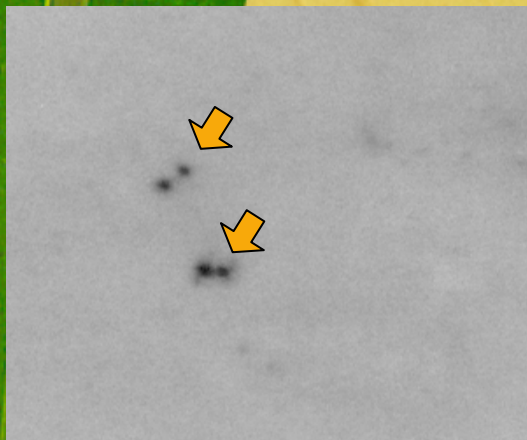
# Musa EST mapping onto Rice Genome



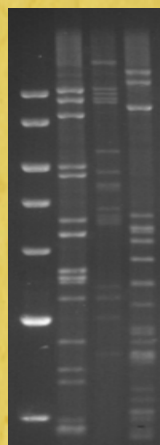
BLASTN (5'-3' EST paired hit ; E value less than  $10^{-20}$ , distances between 5' and 3' EST less than 20kb) identified 192 Musa cDNAs on Rice genome (Build 4 pseudo molecules)

# Banana RGA sequencing

BAC Filter hybridization with selected RGA probes



BAC Fingerprint analysis



*Musa balbisiana*  
Clone PKW (pisang klutuk wulung)

$2n = 2x = 22$

BAC library MBP **Mbp**



*Musa acuminata*  
Clone Calcutta 4

$2n = 2x = 22$

BAC library MA4 **Ma4**

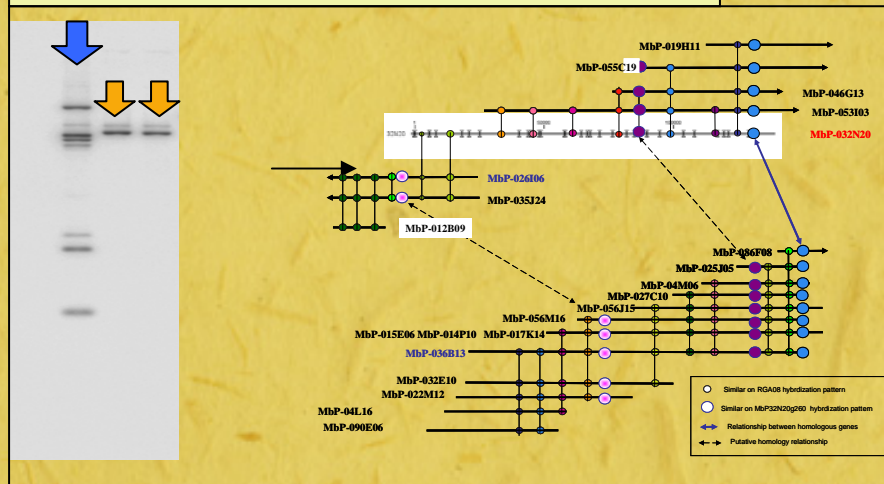


*Musa acuminata*  
Clone Grande Naine

$2n = 3x = 33$

BAC library MAC **MaC**

Southern blot, Hybridization with same selected RGA probes, physical mapping and BAC selection and sequence



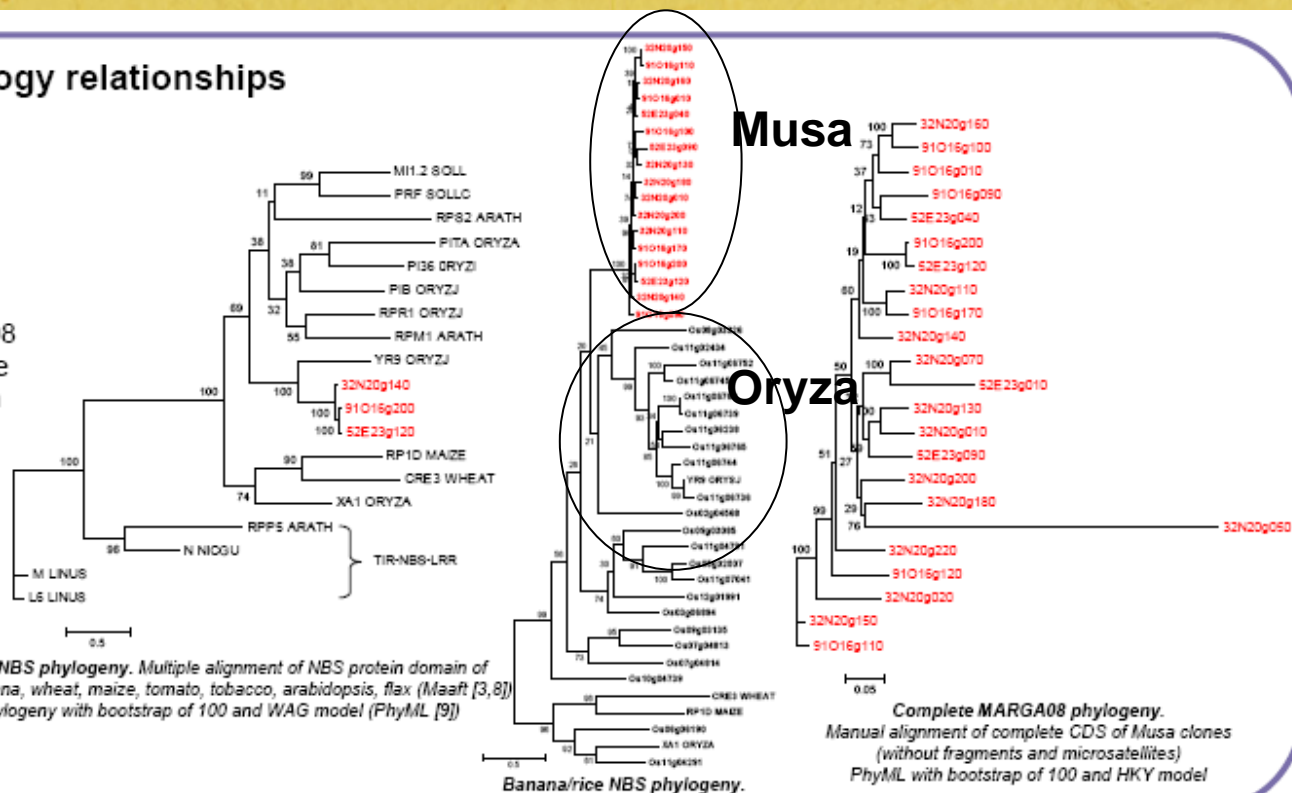
# phylogenetic analysis of resistance gene analogs (RGA)

## Musa/Oryza NSB-LRR homology relationships

- **Plant NBS phylogeny** suggests that we can take RP1D\_MAIZE, CRE3\_WHEAT and XA1\_ORYZA to root the banana/rice NBS phylogeny.

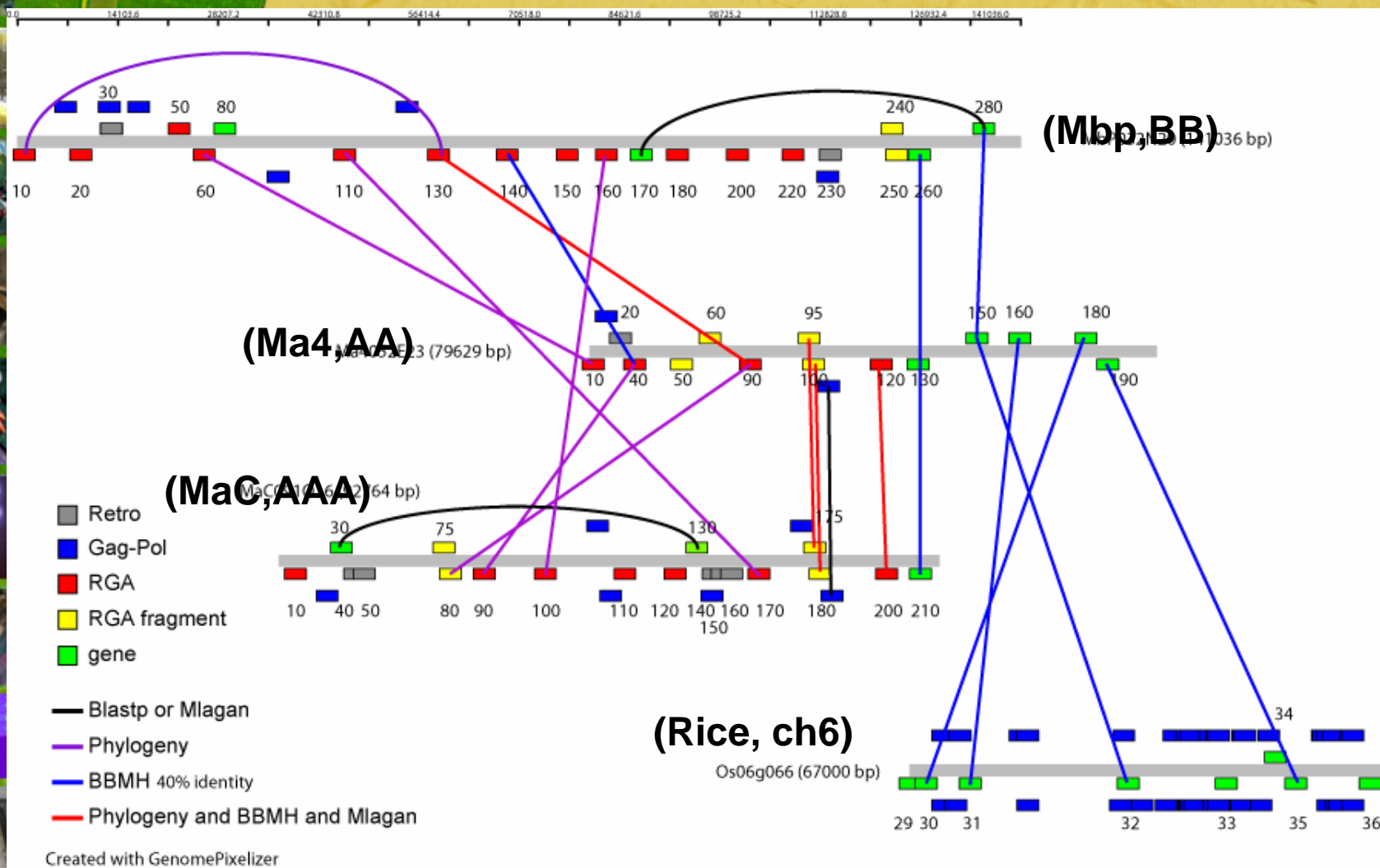
- **Banana/rice NBS phylogeny** suggests the rice chromosome 11 RGA and MARGA08 tandem duplications occurs after banana/rice divergence (MARGA08 cluster younger than Os11RGA cluster).

- **MARGA08 phylogeny** suggests that duplications and mutations leading to pseudogenization or neofunctionalization occurred both before and after Musa speciation.



# Preliminary results of manual annotation:

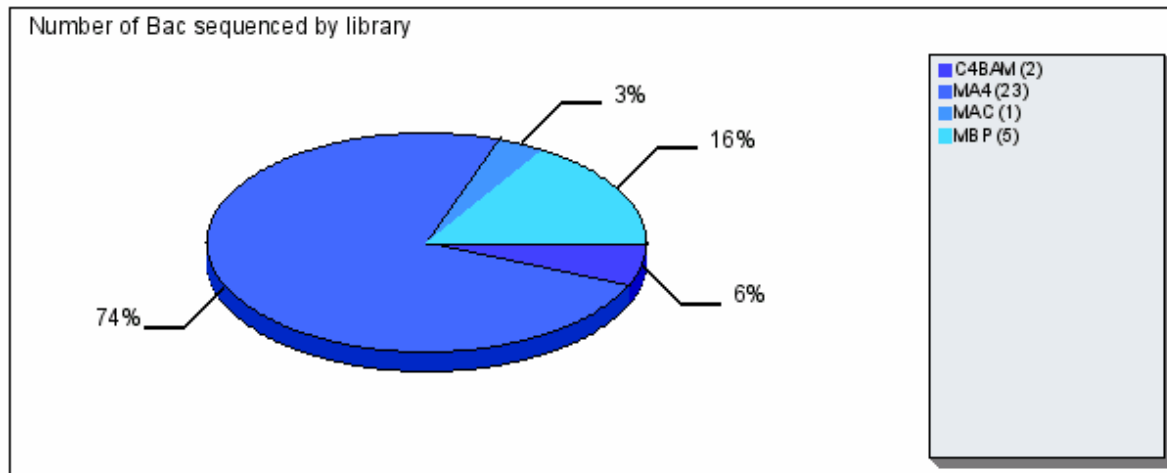
- MARGA08 NBS-LRR type RGA ancestor is monoexonic
- RGA gene duplication occurs before and after A/B speciation



# Statistics and status quo of BAC/BAC-end sequencing

	BAC	BAC END
Number of clones sequenced:	31	6466

**currently 41 BACs**



MBP_91N22	<a href="#">AC186755</a>	BAC	<a href="#">MBP</a>	154.246	Completed	2006-05-28	TIGR	CIR560 ( <a href="#">DQ334869</a> )	B1-3 glucanase gene
MA4_54B05	<a href="#">AC186753</a>	BAC	<a href="#">MA4</a>	54.106	Completed	2006-05-28	TIGR	SB661 ( <a href="#">DQ185893</a> )	thioredoxin
MA4_106O17	<a href="#">AC186747</a>	BAC	<a href="#">MA4</a>	143.796	Completed	2006-05-28	EMBRAPA	SB851 ( <a href="#">DQ185895</a> )	phosphoglycerate kinase
MBP_94I16	-	BAC	<a href="#">MBP</a>	121.63	Completed	-	NIAS	-	-
<a href="#">MA4_86B03</a>	-	BAC	<a href="#">MA4</a>	87.766	Completed	2006-02-12	NIAS	MA4LIMFES002D_F11	WRKY transcription factor gene
MA4_22H07	-	BAC	<a href="#">MA4</a>	-	Characterized	-	NIAS	MUC4LH1002_F07	Class-1 LMW Heat Shock Protein
MA4_25J11	<a href="#">AC186746</a>	BAC	<a href="#">MA4</a>	105.019	Completed	2006-05-28	TIGR	SB132 ( <a href="#">DQ185891</a> )	chlorophyll A-B binding protein type I

# Sequenced BACs by NIAS

Name	Map	Sequence	Size(bp)
A07banana	Map	Sequence	119,244
A39Musa	Map	Sequence	61,335
A40Musa	Map	Sequence	108,948
A41Musa	Map	Sequence	95,303
A42Musa	Map	Sequence	74,174
A43Musa	Map	Sequence	95,242
A44Musa	Map	Sequence	180,124
A45Musa	Map	Sequence	87,766
B45_ADH_Contig1	Map	Sequence	134,662
BMA4-3F3	Map	Sequence	124,825
MA4-2G17	Map	Sequence	37,847
MA4-3G18	Map	Sequence	86,265
MA4-4L11	Map	Sequence	58,704
MA4-9F20	Map	Sequence	92,303
MA4-iJ14	Map	Sequence	104,637
MBP-6L16	Map	Sequence	136,446
MBP-8E04	Map	Sequence	106,121
Musa_B3_125A12	Map	Sequence	110,853
Musa_B46_52E23	Map	Sequence	79,629
Musa_B47_91O06	Map	Sequence	92,764
Musa_B48_32N20	Map	Sequence	141,036
Musa_B4_52N12	Map	Sequence	73,023
Musa_B5_18J06	Map	Sequence	76,327
Musa_B6_12B06	Map	Sequence	91,293

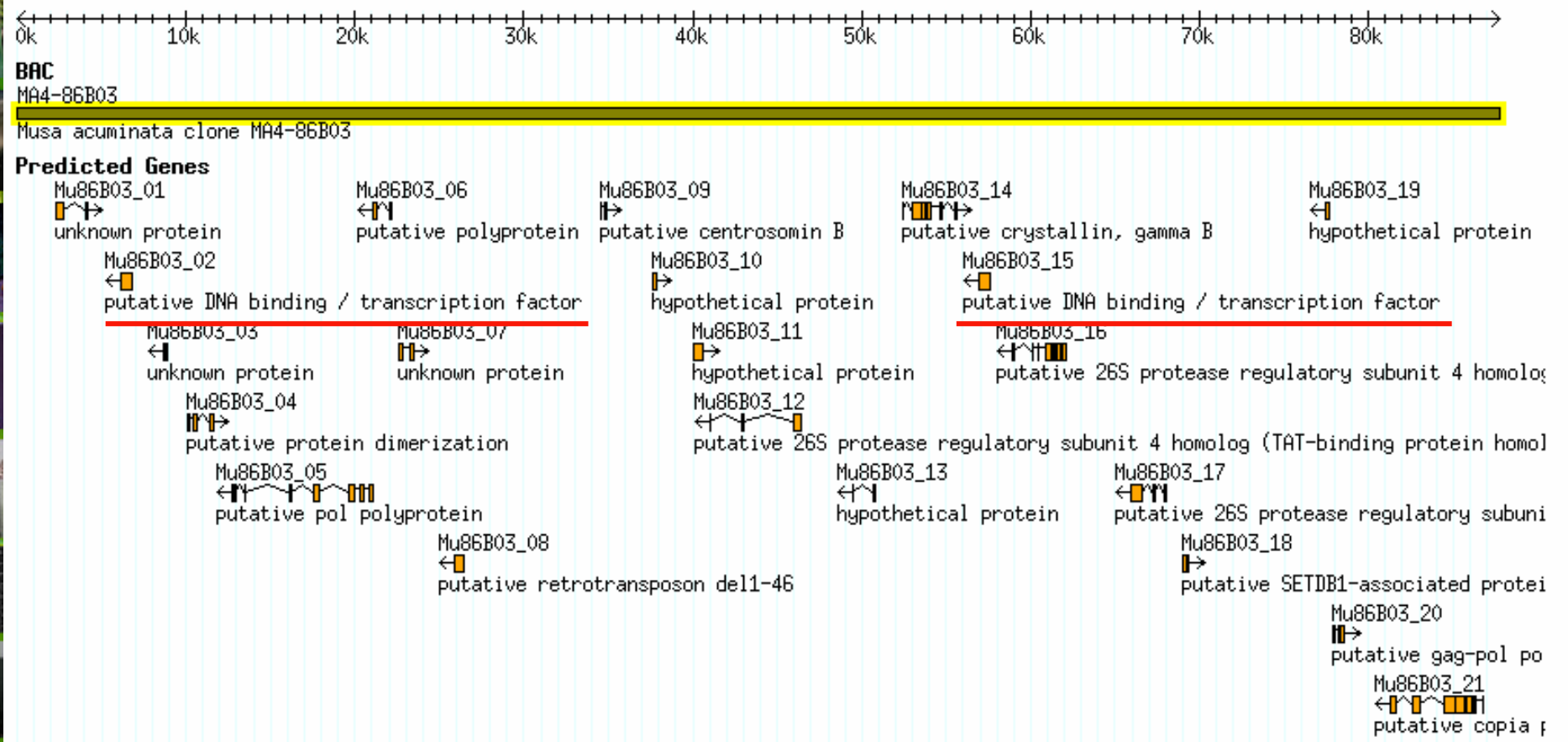
**2,368,871bp**



[www.musagenomics.org](http://www.musagenomics.org)

**BAC: MA4 86B03**  
**probe:WRKY transcription factor gene**

**GBrowse**  
**Physical maps viewer**



# Musa Genome Resources Center (MGRC)

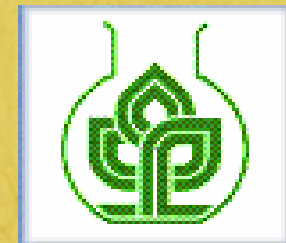
[www.musagenomics.org](http://www.musagenomics.org)

## The mission:

- Acquire
- Store
- Produce
- Distribute Genome resources

## Location:

- IEB, Olomouc, Czech Republic



GMGC Global Musa Genomics Consortium

About GMGC

Genome resources

Research

Projects

News

Contacts

Links

home

Genomic DNA

BAC Libraries

cDNA Libraries

DNA clones

Order

The Musa Genomics Consortium is an international network of investigators communitating Musa (banana) as a model crop for studies of comparative genomics and banana varieties. The consortium currently brings together expertise from 37 institutions in 24 countries. In comparative genomics, Musa is seen as an ideal model for understanding genomic evolution in relation to biotic and abiotic stresses, in a polyploid, vegetatively propagated crop. Members are committed to close collaboration and agree to share materials and resources, including sequence data and enabling technologies. Wherever possible, the products of the Consortium are placed in the public domain and any new varieties developed as a result of the Consortium's work are intended for dissemination to smallholder farmers.

Musa Knowledge Databases



Germplasm Information



Sequencing Status



Genome Browser

### Latest News:

#### GCP ARM Meeting

12-16 September 2007,  
Benoni, South Africa  
[\[more\]](#)

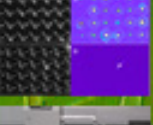
#### Plant Genomics European Meeting

Spain, October  
3-6, 2007  
[\[more\]](#)

#### ISHS/ProMusa Symposium

South Africa, September  
10-14, 2007 [\[more\]](#)

GMGC is an initiative of  
Biodiversity International



# What to do next?

- BAC end sequence the DH Pahang BAC library
- Develop up to 100,000 ESTs
- Develop up to 5000 (SSR and DArT) markers
- Develop F2 population from existing population:  
DH Pahang x P. Madu
- Develop a saturated genetic map
- Anchor physical and genetic map
- Develop a segregation population for drought tolerance identification
- Develop a cytogenetic map
- Increase microsynteny studies with rice Arabidopsis and other models