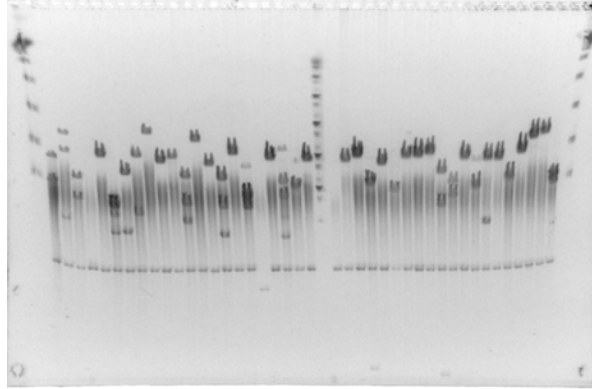


## Synergism with the Cassava Genome Sequencing Consortium

- U. of Arizona, Roche, DOE-JGI sequenced and annotated the AM560-2 genome with 454
- B&MGF project to add transcriptome, genome, and reduced-representation sequencing of additional genotypes

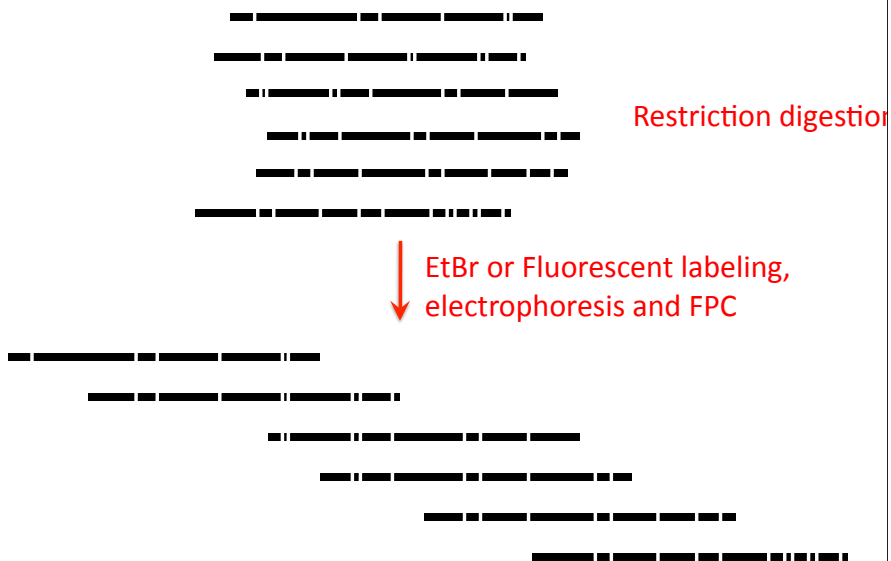
## Fingerprinted AM560-2 BAC library

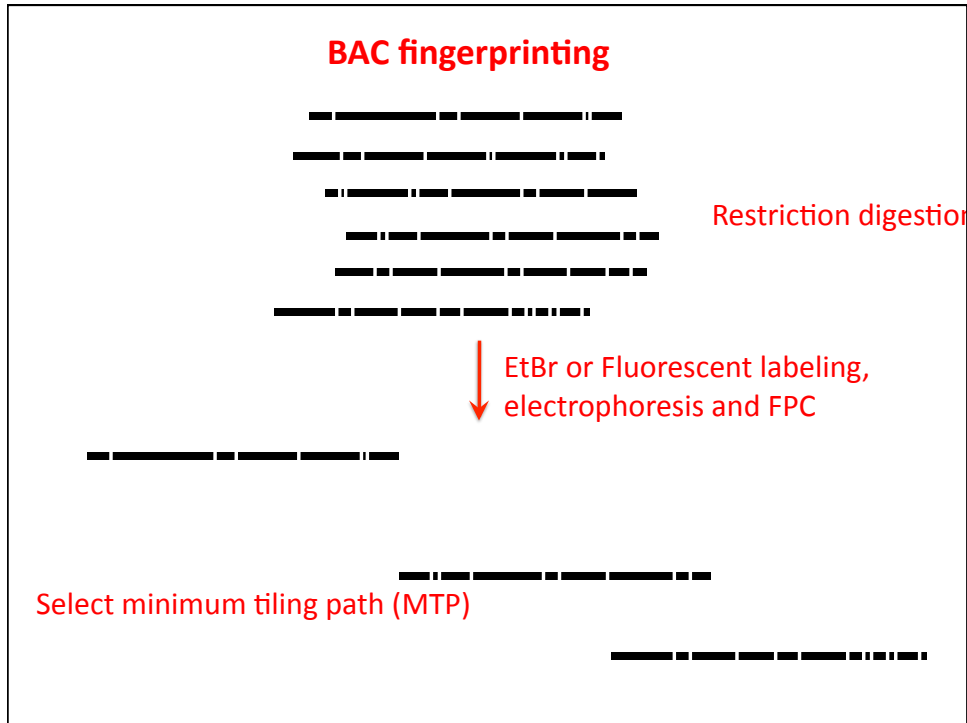
- 72,000 BAC clones, ~10X coverage
- 115Kb average BAC insert size, 7% empty clones



*Library available from CUGI*

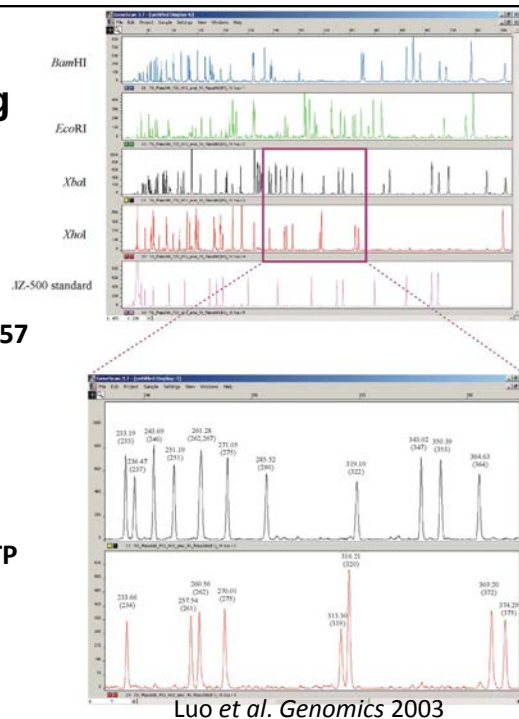
### BAC fingerprinting

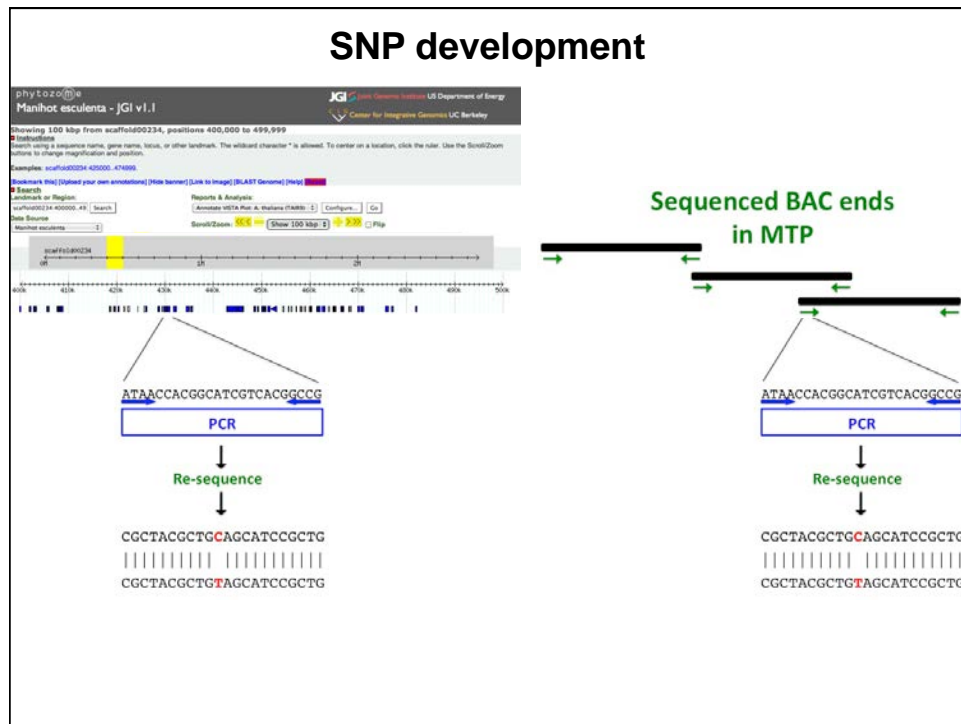




## High information content fingerprinting

- 58,244 BACs assembled
- 2,104 contigs and 5,054 singletons
- Longest contig ~4.8 Mb with 357 clones
- Average contig: 25 BAC clones and 0.41 Mbp
- MTP: 6,868 clones, 710 Mbp
- Sequenced all BAC-ends in MTP and ~2,000 singletons





### SNP development

- 1,536 primer pairs synthesized from BAC-end sequences separated by at least 200 kb
  - 384 PCRs done in 10 cassava genotypes
  - 1,152 PCRs done in the parents of the mapping population
- 1,536 primer pairs synthesized from genes from the cassava genome separated by at least 50 kb
  - All PCRs done in the parents of the mapping population
- 384 SNPs derived from BAC-ends
- 384 SNPs derived from genes
- 48 SNPs derived from ESTs (from Morag Ferguson)

# SNP Discovery

Sequences were assembled and CONSED was used to find SNPs

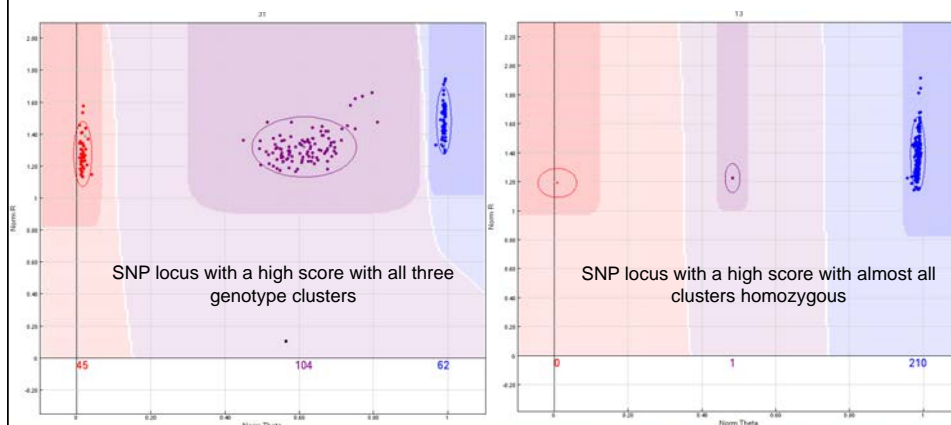
Three good quality SNPs



## Preliminary Illumina genotyping data

- 48 SNPs derived from cassava ESTS in GB
- 39 (81%) out of 48 markers produced scorable results
- 18 (46%) of those were polymorphic in the parents and therefore will be informative in the progeny
- On average, SNPs could be called in 98% of the 216 DNA samples
- 5 (2.3%) progeny appear to be off-type (do not carry either of the parental alleles)
- 3 SNP genotype calls out of 10368 (0.03%) appear to be scoring errors
- 3 DNA samples (1.4%) failed for all assays

## Examples of High Quality SNP Genotyping Clusters for the Mapping Population (48-plex)



## Capacity building

- A SNP genotyping workshop held in November 9-13, 2009 at the University of Pretoria, South Africa
- Participants brought DNA from multiple cassava genotypes
- SNP diversity in cassava

## Coming up in the Near Future

- Genotype the mapping population for 384 SNPs derived from the physical map
- Genotype the mapping population for 384 SNPs derived from the genes annotated in the genome
- Genotype all cassava genotypes available for genetic diversity study
- Build genetic map and anchor to the physical map
- Integrate all data with the UA's B&MGF and JGI/Roche cassava genome project
- Integrate with drought phenotype data for QTL mapping

## ***Acknowledgements***

### **IGS-UMB, MD, USA:**

Pablo Rabinowicz  
Kimm O'Brien  
Kevin Galens  
Maximo Rivarola

### **ACGT, South Africa:**

Zander Myburg, University of Pretoria  
Jane Morris, University of Pretoria  
Chris Rey, University of the Witwatersrand

### **Danforth Center, St. Louis, MO:**

Martin Fregene

### **UC Davis, CA:**

Mingcheng Luo  
Yaqin Ma

### **IITA, Nairobi, Kenya:**

Morag Ferguson

### **CIAT, Cali, Colombia:**

Hernan Ceballos  
Augusto Becerra Lopez-Lavalle

**National Root Crops  
Research Institute, Nigeria:**  
Emmanuel Okogbenin

### **Funding:**



**BILL & MELINDA  
GATES foundation**