Advanced Genomic Technologies for Genetic Enhancement of Yam (*Dioscorea* spp.)

Ranjana Bhattacharjee, Antonio Lopez-Montes, Michael Abberton, P. Lava Kumar, and Robert Asiedu

r.bhattacharjee@cgiar.org

*International Institute of Tropical Agriculture (www.iita.org)*

*Sweet Potato and Yam Session*

19 Jan 2016

*World Congress on Root and Tuber Crops, Nanning, China*
Challenges with clonally propagated crops

- Long breeding cycle
- Low multiplication ratio of planting materials
- Vegetative propagation
- Heterozygous genetic background
- Polyploidy (in yams)
- Poor to no flowering/synchronous flowering
- Dioecious (yams)
- Mislabelling/presence of duplicates
- Labor-intensive
- Little or poor knowledge and information available on useful traits for use in improvement programs
- Low genetic diversity in farmers’ field
- Biotic stresses (mainly fungi and viruses)
Facts/Opportunities: Yam

**Different species:** >600 species, of which 10 are cultivated

- D. alata
- D. dumetorum
- D. bulbifera
- D. rotundata

**Variability within each species:** Opportunities for crop improvement

- D. rotundata
- D. alata

- Food security
- Resilience
- Sustainability
- Income Generation
- Hybrids/Varieties

A member of CGIAR consortium | www.iita.org
Various targets for crop improvement

- Yield potential and yield stability
  - Photosynthesis efficiency
  - Harvest index
  - Reduced inputs (fertilizers, pesticides, etc.)

- Adaptation to climate change
  - Tolerance (drought, heat, etc.)
  - Avoidance
  - Post stress recovery

- Durable resistance to biotic stress
  - Existing pests and diseases (virus, fungi, nematodes)
  - New pests and diseases
  - Invasive species

- Quality and value-added products
  - Starch quality and quantity
  - Consumer preference
  - Food safety aspects
Completing the Whole Genome Sequencing of six important cultivated *Dioscorea* spp.
- *D. rotundata* (IITA-IBRC-JIRCAS)
- *D. alata* (IITA-TGAC)
- *D. dumetorum* (underway, IITA-AOCC)
- *D. cayenensis, D. esculenta, D. bulbifera*

Development of genomic resources

Development of mapping populations for different target traits

Development of training populations

Genotyping-by-Sequencing:
- understanding population structure, genetic diversity, varietal identification,
- linkage mapping and identification of QTLs
- GWAS and GS

Metabolomics and high-throughput phenotyping

Application of transcriptomics

Tissue culture and Genetic transformation
### Whole Genome Sequencing

#### Basic assembly stats

- **No. contigs**: 57,706
- **Largest contig**: 296.5 kb
- **Total length**: 620 Mb
- **GC (%)**: 36.05
- **N50**: 19.3 kb
- **No. gene models**: 40,055

- 91.25% properly paired reads mapping
- 96.84% of 44,134 ncbi *D. alata* ESTs hit with BLAST
- Over 90% of Core BUSCO Plantae genes present: C:88%[D:25%], F 6.0%, M:5.5%, n:956
Yam: GCDT project

Objective: True-to-type/duplicate identification in national and international germplasm collection; establishment of a global DNA bank

- Priority species: *Dioscorea alata*, *D. rotundata*, and *D. cayanensis*
- Number of accessions:
  - IITA: 1500 (*alata* = 815; *cayanensis* = 59; *rotundata* = 626)
  - NARS: 785 (mostly *D. alata*)
  - Countries: Benin, Ghana, Togo, Philippines, Costa Rica, New Caledonia, Papua New Guinea, Fiji and Thailand (No samples received from Cote d’Ivoire, Solomon Island, Vanuatu and Vietnam)
- Genotyping with 50 SSR (18 genomic and 32 EST-SSR) markers

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated genome size</td>
<td>1Gb</td>
</tr>
<tr>
<td>Number of samples</td>
<td>96</td>
</tr>
<tr>
<td>Index depth</td>
<td>48 individuals per lane</td>
</tr>
<tr>
<td>Phenotype</td>
<td>Anthracnose disease trait</td>
</tr>
<tr>
<td>Sequencing type</td>
<td>Illumina HiSeq 2500 (1-2 lanes)</td>
</tr>
<tr>
<td>Read type</td>
<td>1x100bp Single end</td>
</tr>
</tbody>
</table>

Averages from 96 Samples [using TASSEL (UNEAK) pipeline]
- Unique Tag Counts: 83,018
- Raw Reads: 3,156,223
- Aligned Reads: 2,535,879
- Coverage per unique tag: 30

Figure. Shared SNP distribution contained in at least 50% of the individuals
GBS-based linkage mapping

Figure 4. Genetic map for parent 1 and 2.


**Association analysis**

**Association analysis for anthracnose disease with 10,077 SNPs**

Table. Disease state and phenotypic scores for anthracnose disease

<table>
<thead>
<tr>
<th>Anthracnose rating scale</th>
<th>Range</th>
<th>Meaning</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 - 2% plant area affected with anthracnose</td>
<td>Plant healthy or with a trace of disease</td>
<td>Highly resistant</td>
</tr>
<tr>
<td>2</td>
<td>&gt;2 - 10% of plant area with symptom of anthracnose</td>
<td>Plant healthy, with more observable anthracnose</td>
<td>Resistant</td>
</tr>
<tr>
<td>3</td>
<td>&gt;10 - 25% of plant area with anthracnose symptoms</td>
<td>Plant unhealthy, with conspicuous anthracnose</td>
<td>Moderately resistant</td>
</tr>
<tr>
<td>4</td>
<td>&gt;25 - 50% of plant area diseased</td>
<td>Plant unhealthy, with large anthracnose lesions</td>
<td>Susceptible</td>
</tr>
<tr>
<td>5</td>
<td>&gt;50% of plant area affected by anthracnose</td>
<td>Lesions coalesced, plant dead</td>
<td>Highly susceptible</td>
</tr>
</tbody>
</table>

Table. Significant SNP loci from the 705 filter subset for both test 1 and 2.

<table>
<thead>
<tr>
<th>Association_test</th>
<th>Locus</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test_1_705</td>
<td>TP6683</td>
<td>0.003</td>
</tr>
<tr>
<td>Test_1_705</td>
<td>TP22</td>
<td>0.004</td>
</tr>
<tr>
<td>Test_1_705</td>
<td>TP7189</td>
<td>0.004</td>
</tr>
<tr>
<td>Test_1_705</td>
<td>TP6186</td>
<td>0.007</td>
</tr>
<tr>
<td>Test_1_705</td>
<td>TP6978</td>
<td>0.007</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP11455</td>
<td>0.002</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP7744</td>
<td>0.003</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP6683</td>
<td>0.006</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP12241</td>
<td>0.007</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP760</td>
<td>0.008</td>
</tr>
<tr>
<td>Test_2_705</td>
<td>TP12061</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Test 1: Tails of disease score rating scale (1-2) and (4-5) = 155 SNPs p-value < 0.01

Test 2: Disease score rating scale (1-2.9) and (3-5) = 424 SNPs p-value < 0.01

Only 45% of GBS profiles could be annotated to de novo sequencing data from parents.
Focus: Genotyping by sequencing (GBS); high-throughput phenotyping (metabolomics); morphological characterization; breeding applications (inter- and intra-specific crosses)

- GBS of 810 *D. rotundata* genotypes (core collection = 470 landraces; breeding lines = 307 genotypes; varieties from markets = 33) completed
- Different bioinformatics pipeline tested: a high proportion of filtered reads aligned to the reference *D. rotundata* genome using bowtie 2 software
- A customized R script identified 3068 polymorphic loci of which 55.9% were bi-allelic and 44.1% multi-allelic.
- Morphological characterization (above and under ground traits) of 810 genotypes (in three replications: head, middle and tail portions) completed for two years

Genotyping by sequencing (GBS) reveals the complex genetics of *D. rotundata*. This sequence alignment shows the heterozygous sequence of the reference genome and three alleles found in *D. rotundata* germplasm. The number of alleles per sample ranged between 1 and 3, indicating that some clones are polyploids.

Total reads processed: 438,685,435
Reads with adapters: 340,851,682 (77.7%)
Reads written (passing filters): 438,685,435 (100.0%)
Total basepairs processed: 44,307,228,935 bp
Quality-trimmed: 2,290,709,390 bp (5.2%)
Total written (filtered): 25,613,093,559 bp (57.8%)
Shorter than 64bp: 273,540,127 (62.4%)
Not assigned to barcode: 12,892,837 (2.9%)
Total reads passing all filters: 152,252,471 (34.7%)
Phenotyping of GBS materials

- All 810 genotypes planted in an augmented design using three checks
- Each tuber was cut into three sections (head, middle and tail) for planting
- Planting in two locations including Ibadan and Ikenne, in May and June, 2014 respectively
- Observations on following above ground traits such as:
  - Days for germination (Earliness)
  - Number of vines per portion of the tuber
  - Pests and diseases (anthracnose, virus, and nematodes)
  - Leaf shape and number of internodes
  - Flowering traits (monoecy, dioecy, no flowering)
- For below ground observations:
  - Tuber number and shape
  - Tuber weight
- Soil samples from around each tuber and sent for analysis
49 accessions from IITA breeding program
5 species (D. alata, D. bulbifera, D. cayenensis, D. dumetorum, D. rotundata)
Tuber material (Head, middle and tail; selected leaf material)

- No significant difference in metabolite composition between tuber and leaf sampling
- Slight gradation in metabolite composition across different portions
- Only few metabolites correlate in pathways: due to postharvest deterioration
### Inter- and Intra-specific Crosses

<table>
<thead>
<tr>
<th>Trait of Interest</th>
<th>No. of female</th>
<th>No. of male</th>
<th>Crosses</th>
<th>Progenies</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUE</td>
<td>6</td>
<td>6</td>
<td>TDr 04-219</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 95/01932</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 89/02157</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 89/02665</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 97/00777</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 97/00917</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDr 89/02157</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDa 02/00012</td>
<td>341</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TDa 98/01166</td>
<td>4</td>
</tr>
</tbody>
</table>

Conventional karyotyping of *Dioscorea* spp.
IITA-CIRAD-INRA-CTCRI project: Assessment of genetic diversity in *D. alata* germplasm

Student: Ph.D. student (Claudie Pavis), France

- 384 *D. alata* germplasm genotyped with 34 SSR markers
- A total of 847 alleles recorded across 26 SSRs with mean number of alleles ranging from 10.6 (CIRAD) to 7.5 (IITA)


<table>
<thead>
<tr>
<th>Indicators</th>
<th>INRA</th>
<th>CIRAD</th>
<th>CTCRI</th>
<th>IITA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>129</td>
<td>83</td>
<td>82</td>
<td>90</td>
</tr>
<tr>
<td>A&lt;sup&gt;t&lt;/sup&gt;</td>
<td>217</td>
<td>255</td>
<td>194</td>
<td>181</td>
</tr>
<tr>
<td>A&lt;sup&gt;m&lt;/sup&gt;</td>
<td>9.0</td>
<td><strong>10.6</strong></td>
<td>8.1</td>
<td>7.5</td>
</tr>
<tr>
<td>A&lt;sup&gt;s&lt;/sup&gt;</td>
<td>10</td>
<td>48</td>
<td>14</td>
<td>10</td>
</tr>
</tbody>
</table>
Genetic Transformation of Yam

- **Agrobacterium**-mediated transformation system established for yam using apical meristems.
- Transgenic plants generated using reporter genes and validated by molecular analysis.
- Significant difference in transformation efficiency was observed among different cultivars transformed.
- It takes 5–6 months from transformation to regeneration of complete transgenic plant.
- The transformation protocol is validated with different accessions of *D. rotunda* and *D. alata*.

Agrobacterium mediated transformation of yam using embryogenic callus

- Regeneration through somatic embryogenesis of yam has been established at IITA-Nairobi.
- Embryogenic calli were transformed with Agrobacterium using gusA reporter gene.
- Transformed calli are currently under selection and regeneration.

Poster No. P0406
Completing the reference genome sequencing of remaining *Dioscorea* spp. mainly *D. cayenensis*, *D. dumetorum*, *D. esculenta* and *D. bulbifera*

- Development of species-specific and cross-species markers
- Understanding the complex traits such as flowering, sex determination, diseases, and quality traits
- NSF-BREAD project: PacBio sequencing and re-sequencing of potential parents
- Understanding nematode resistance using transformation
- Understanding the complexities with flowering traits
- Use of genomics-assisted information in breeding
Partnerships

- RTB team of CGIAR CRP
- Cornell University, USA
- Royal Holloway University of London, UK
- The Genome Analysis Center, UK
- African Orphan Crops Consortium
- Iwate Biotechnology Research Center, Japan
- The Institute of Experimental Biology, Czech Republic
- JIRCAS, Japan
- CIRAD, France and Guedeloupe
- UC Berkeley, USA
- Michigan State University, USA
- National programs within Africa, Asia and the Pacific
- Universities: Masters and PhD students