Application of Genetic Fingerprinting Technique to Track Diffusion and Adoption of Improved Rice and Maize Varieties in Tanzania

Leonard Oruko, Diiro Gracious, Fred Tairo and Sam Amanquah
Background and Objectives

RATIONALE FOR USE OF DNA FINGERPRINTING
Why Genetic Fingerprinting

Traditional approach to diffusion and adoption estimates relies on data from
- Farmer/respondent recall
- Expert opinion of breeders and seed value chain actors
- Records

The DIIVA Study employed all the three and revealed the following
- Farmers face challenges with identification of varieties in their fields
- In some cases, adoption estimates were inconsistent with previous studies

LSMS-ISA has revealed some degree of systematic measurement error from HH surveys in a number of African countries
- Cropped area estimates
- Output/yield estimates

The Maize and wheat study in Ethiopia
Objectives

- Test the logistical and technical feasibility of estimating adoption using both farmer call information and genetic fingerprinting information
- Provide more precise estimates of adoption of improved rice and maize varieties in the pilot study areas
- Provide an estimate of adoption of improved crop varieties whose release were supported by the AGRA Program for Africa’s Seed Systems (PASS)
- Develop the capacity of the Department of Research and Development (DRD) to conduct a similar study in Tanzania at scale
Key Learning Questions

Cost of DNA fingerprinting has dropped significantly over the years. It is therefore feasible to obtain more precise estimates of \textit{varietal diffusion and adoption levels} through \textit{DNA fingerprinting (the Gold Standard???)}

1. Test the technical feasibility: Can DNA fingerprinting discriminate across improved varieties grown by farmers?
   - Rice and Maize grain samples from farmer fields
   - Reference library of breeder seed

2. Test the logistical feasibility: What is the optimal mix of implementing organizations and institutional arrangements?
   - Can the organizations form functional partnerships for the collection of data and grain samples while preserving identity; conduct analysis and generate credible results?
Approach and Method

COMBINING TRADITIONAL HOUSEHOLD SURVEYS WITH DNA ANALYSIS
Approach: Institutional arrangements and workflow process

1. Filled questionnaires
2. Test samples harvested crop from a 4x4 area
3. Similar code for grain bags and questionnaires

ORUKO Inc! Analysis and Report

REPOA

Rice and maize breeders from DRD

1. DNA extraction from test and reference materials
2. DNA quality and quantity tests & documentation
3. DNA shipment to DArT

Mikocheni Biotech Lab

DNA received by DArT for analysis
Approach: Sampling and data collection

- **Superimpose on the AGRA outcome panel survey**
  - Kilosa, Kilolo, Mbarali and Sumbawanga selected
    - Planned to collect rice and maize samples and administer a short duration questionnaire to 725 farm households
    - Iringa Rural re-introduced to increase target to 800 households

- Enumerators trained on grain sample collection and use of ODK platform for data collection

- Reference material comprising released improved rice and maize varieties collected as well
Data and sample collection

- Grain sampling protocol entailed collection of 10 maize cobs and 30 panicles of rice per field
- HH socio-economic data capture thereafter
- Sample and data collection commenced after some farmers had harvested
- Maize and rice farmers were replaced
- Data and test samples collected from rice farmers storing harvest from a given field in identifiable bags
Grain sample preparation

- Reference and test samples were packaged labeled and sent to MARI
- The Grain was Grinded into fine flour packed into tubes and shipped to BeCA labs for DNA extraction
DNA extraction

Two technicians from MARI conducted the DNA extraction at BeCA labs in Nairobi

High quality DNA was extracted following the protocol for Bulk Seed DNA Extraction (few rice samples showed degradation)

The purity levels were relatively high ranging from 1.79 to 1.93 for maize and 1.58 to 2.02 for rice

DNA concentrations for both rice and maize reference material were sufficient for downstream applications
DNA Analysis: Simplified View of DArT Genetic ID Assay Process

(1). Seed Sample (>1000)

(2). DNA Extracted from flour

(3). Assay for thousands of markers
   - (a). DArTseq
   - (b). DArT Array

(4). Reference Development

(5). Testing Result
   - Clean “A”
   - “B” Contaminated with “A”

[Charts and graphs showing quantitation of markers A, B, C, etc.]
Findings

DIFFUSION AND ADOPTION
Summary statistics

Certified maize seed from the seed companies were collected instead of breeder seed

- A total of 71 varieties were collected for the development of maize reference library
- 15 rice varieties in the Official Varieties Release List from TOSCI and 3 varieties under the NPT

<table>
<thead>
<tr>
<th>Crop/Region</th>
<th>Region/district</th>
<th>2013 panel survey</th>
<th>2015 survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize</td>
<td>Iringa rural district</td>
<td>189</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Kilolo district</td>
<td>206</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>395</strong></td>
<td><strong>291</strong></td>
</tr>
<tr>
<td>Rice</td>
<td>Morogoro</td>
<td>238</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Iringa</td>
<td>17</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Mbeya</td>
<td>290</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>Rukwa</td>
<td>274</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>819</strong></td>
<td><strong>356</strong></td>
</tr>
</tbody>
</table>
Diffusion of improved varieties-farmer recall

**Reported use of improved maize varieties (n=244)**

- **Iringa vijiini**
- **Kilolo**
- **Total Iringa Region**

**Reported use of improved rice varieties (n=328)**

- **All regions**
- **Morogoro**
- **Iringa**
- **Mbeya**
- **Rukwa**

**Criteria for farmer identification of improved maize varieties n=133)**

- **Certified seed**
- **Seed non-certified but from reliable source**
- **Not sure**

- Only 45% of maize farmers planted improved varieties in 2015
- Only 6% of rice farmers planted improved varieties in 2015
Maize Varietal Identification – Farmer Recall n=291
Rice Varietal Identification—Farmer Recall n=365
The Identified Primary Constituent (IPC)

- The purity level of the IPC expressed as a percentage was the main criteria for variety identification.
- IPC purity level of 70% adopted as the minimum threshold for correctly identifying a variety.
- The entire rice test sample had 99% and above IPC purity levels.
- A few of the submitted maize test samples had less 70% IPC purity.

<table>
<thead>
<tr>
<th>Region /district</th>
<th>70% and above IPC purity level</th>
<th>Less than 70% IPC purity level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of samples</td>
<td>%</td>
</tr>
<tr>
<td>Kilolo district (n=185)</td>
<td>174</td>
<td>94.1</td>
</tr>
<tr>
<td>Iringa rural district (n=138)</td>
<td>136</td>
<td>98.6</td>
</tr>
<tr>
<td>Iringa region (n=323)</td>
<td>310</td>
<td>95.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Crop</th>
<th>Reference samples</th>
<th>Test Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Customer genotype</td>
<td>IPC</td>
</tr>
<tr>
<td>Maize</td>
<td>H 625 Kitale</td>
<td>H625 kitale</td>
</tr>
<tr>
<td></td>
<td>Katumani</td>
<td>Katumani</td>
</tr>
<tr>
<td>Rice</td>
<td>Tai (ir 0334262)</td>
<td>Tai (ir 0334262)</td>
</tr>
<tr>
<td></td>
<td>Supa India</td>
<td>Supa India</td>
</tr>
</tbody>
</table>
Maize Varietal Identification - DNA Fingerprinting

- Misclassification of varieties from farmer recall
- Most test materials can be classified as improved
- Kitale hybrids most popular as opposed to “Kienyeji”
## Rice Varietal Identification - DNA Fingerprinting

<table>
<thead>
<tr>
<th>Region</th>
<th>SUPA INDIA</th>
<th>TXD 306</th>
<th>TXD 85</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>All regions (n=444)</td>
<td>430</td>
<td>96.85</td>
<td>12</td>
</tr>
<tr>
<td>Iringa (n=48)</td>
<td>46</td>
<td>95.83</td>
<td>2</td>
</tr>
<tr>
<td>Mbeya (n=219)</td>
<td>200</td>
<td>91.32</td>
<td>9</td>
</tr>
<tr>
<td>Morogoro (n=72)</td>
<td>72</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Rukwa (n=103)</td>
<td>101</td>
<td>98.06</td>
<td>1</td>
</tr>
</tbody>
</table>

- Misclassification of varieties from farmer recall
- All of the submitted test material can be classified as improved
- Supa India most popular as opposed to Zambia
**Degree of misclassification**

- Even after relaxing the 70% threshold, 82% of maize samples were misclassified.

- Only 23% of rice samples collected were correctly classified.

---

**Degree of misclassification of maize varieties**

- **n=444** Total sample, 48 Iringa, 219 Mbeya, 72 Morogoro, 103 Rukwa

**Degree of misclassification of rice varieties**

**n=444** Total sample, 48 Iringa, 219 Mbeya, 72 Morogoro, 103 Rukwa
Source of planting material

Reported source of maize seed

Reported Source of rice seed
Seed Management Practices

Maize farmers

- Frequency:
  - Does not change
  - Every year
  - Every 2 years
  - Every 3 years

- Percent of farmers:
  - All households
  - Adopters
  - Non-adopters

Rice farmers

- Frequency:
  - Does not change
  - Every year (2 years)
  - Every 3 years
  - Every 4 years
  - Every 5 years

- Percent of farmers:
  - All households
  - Adopters
  - Non-adopters
Misclassification - High Level Observations

Farmer recall information under-estimates the levels of varietal diffusion and adoption

Misclassification observed
- Does this obtain in the case of other inputs / measurement error?

Farmers simply do not know what seeds they are acquiring - maize

Are traders mis-labeling seed packets / adulteration?

Local / colloquial v. 'formal' variety names

Interpretation / perception of what is an 'improved' variety