AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MSU International Development Working Paper

Testing Alternative Methods of Varietal Identification Using DNA Fingerprinting: Results of Pilot Studies in Ghana and Zambia

by

Mywish K. Maredia, Byron A. Reyes, Joseph Manu-Aduening, Awere Dankyi, Petan Hamazakaza, Kennedy Muimui, Ismail Rabbi, Peter Kulakow, Elizabeth Parkes, Tahirou Abdoulaye, Enid Katungi, and Bodo Raatz

Department of Agricultural, Food, and Resource Economics
Department of Economics
MICHIGAN STATE UNIVERSITY
East Lansing, Michigan 48824

MSU International Development Working Paper No. 149
October 2016

MSU Agricultural, Food, and Resource Economics Web Site: http://www.afre.msu.edu/
MSU Food Security Group Web Site: http://fsg.afre.msu.edu/index.htm

MSU is an affirmative action, equal opportunity employer

MSU INTERNATIONAL DEVELOPMENT PAPERS

The Michigan State University (MSU) International Development Paper series is designed to further the comparative analysis of international development activities in Africa, Latin America, Asia, and the Near East. The papers report research findings on historical, as well as contemporary, international development problems. The series includes papers on a wide range of topics, such as alternative rural development strategies; nonfarm employment and small scale industry; housing and construction; farming and marketing systems; food and nutrition policy analysis; economics of rice production in West Africa; technological change, employment, and income distribution; computer techniques for farm and marketing surveys; farming systems and food security research.

The papers are aimed at teachers, researchers, policy makers, donor agencies, and international development practitioners. Selected papers will be translated into French, Spanish, or other languages.

Copies of all MSU International Development Papers, Working Papers, and Policy Syntheses are freely downloadable in pdf format from the following Web sites:

MSU International Development Papers
http://fsg.afre.msu.edu/papers/idp.htm
http://ideas.repec.org/s/ags/mididp.html

MSU International Development Working Papers
http://fsg.afre.msu.edu/papers/idwp.htm
http://ideas.repec.org/s/ags/midiwp.html

MSU International Development Policy Syntheses
http://fsg.afre.msu.edu/psynindx.htm
http://ideas.repec.org/s/ags/midips.html

Copies of all MSU International Development publications are also submitted to the USAID Development Experience Clearing House (DEC) at: http://dec.usaid.gov/

TESTING ALTERNATIVE METHODS OF VARIETAL IDENTIFICATION USING DNA FINGERPRINTING: RESULTS OF PILOT STUDIES IN GHANA AND ZAMBIA

by

Mywish K. Maredia, Byron A. Reyes, Joseph Manu-Aduening, Awere Dankyi, Petan Hamazakaza, Kennedy Muimui, Ismail Rabbi, Peter Kulakow, Elizabeth Parkes, Tahirou Abdoulaye, Enid Katungi, and Bodo Raatz

October 2016

Maredia is Professor, International Development, Michigan State University; Reyes is Scientist in the Decision and Policy Analysis unit, Katungi is Scientist in the bean research program, and Raatz is Bean Breeder and Geneticist at the International Center for Tropical Agriculture (CIAT); Manu-Aduening is Cassava Breeder at the Council for Scientific and Industrial Research-Crops Research Institute (CRI), Ghana; Dankyi is Agricultural and Socio Economist of Agricultural Innovation Consult., Ghana; Hamazakaza is Agricultural Economist and Muimui is Bean Breeder/Program Leader at the Zambia Agricultural Research Institute (ZARI); and Kulakow is Cassava Breeder/Geneticist, Rabbi is Cassava Molecular Geneticist, Parkes is Harvest Plus Cassava Breeder, and Abdoulaye is Agricultural Economist at the International Institute of Tropical Agriculture (IITA).

ISSN 0731-3483

© All rights reserved by Michigan State University, 2016.

Michigan State University agrees to and does hereby grant to the United States Government a royalty-free, non-exclusive and irrevocable license throughout the world to use, duplicate, disclose, or dispose of this publication in any manner and for any purposes and to permit others to do so.

Published by the Department of Agricultural, Food, and Resource Economics and the Department of Economics, Michigan State University, East Lansing, Michigan 48824-1039, U.S.A.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Standing Panel on Impact Assessment of the Independent Science and Partnership Council under the Strengthening Impact Assessment in the Consultative Group on International Agricultural Research (CGIAR) program funded by the Bill and Melinda Gates Foundation and other donors. We would like to thank all the farmers and survey staff who participated in the data collection efforts for this study in Ghana and Zambia. We also acknowledge field and lab support provided by Gezahegn Girma, Clare Mukankusi, Allan Male and Ramu Punna. Financial support from the CGIAR Research Program on Roots, Tubers and Bananas and PABRA towards some of the field work is duly acknowledged. The authors also acknowledge the editing and formatting assistance of Patricia Johannes.

All opinions and errors in data interpretation and reporting remain with the authors.

EXECUTIVE SUMMARY

Varietal adoption based on household surveys has mostly relied on farmers' response to varietal identification. This method can give biased estimates if farmers are unable to identify improved varieties as a group or by name, or give names that do not match with the released variety list. To tackle these potential problems other innovative methods have been suggested that require time and resource intensive data collection such as including follow-up questions in the survey instrument to gather information on varietal traits, visiting the field to observe plant characteristics, or collecting sample materials (i.e., photos, seeds/plant tissues) from the farmers for later verification by experts. However, the accuracy of these different methods for identifying varieties grown by farmers to be able to estimate variety specific adoption is unknown.

This paper reports the results of two pilot studies conducted in Ghana and Zambia to test and validate some of these different approaches of collecting variety-specific adoption data against the benchmark of DNA-fingerprinting to determine which method can accurately identify released varieties used by farmers. Results suggest large variations in the estimates of varietal adoption obtained by these different methods compared to DNA fingerprinting results. Results also point to potential challenges of these alternative methods of varietal identification, including DNA fingerprinting in a developing country setting. The implications of these results on future adoption and impact studies are discussed.

CONTENTS

ACKNOWLEDGMENTS iii
EXECUTIVE SUMMARY iv
LIST OF TABLES vi
ACRONYMS viii
1. INTRODUCTION 1
2. METHOD AND APPROACH 3
2.1. Methods of Varietal Identification Evaluated 3
2.2. DNA Fingerprinting 4
2.3. Sampling and Data Collection 5
2.3.1. Ghana 5
2.3.2. Zambia 5
2.4. Varietal Identification Using Expert Elicitation Based on Photographs and Seed Samples 7
3. RESULTS 8
3.1. Classification of Reference Library Materials and Farmer Samples into Variety Groups Based on DNA Fingerprinting 8
3.2. Classification of Varieties by Farmers' Reported Name and Type of Variety (Method A) 12
3.3. Classification of Varieties by Farmer Reported Characteristics of the Plants (Cassava Only) (Method B) 14
3.4. Classification of Varieties by Farmers' Based on Matching their Variety with Seed Samples Shown (Beans Only) (Method C) 15
3.5. Classification of Varieties by Cassava Experts Visiting the Field and Recording Observed Morphological Traits (Cassava Only) (Method D) 16
3.6. Classification of Varieties by Experts Looking at Pictures of Cassava Plants and Pictures of Bean Seeds (Method E) 18
3.7. Classification of Varieties by Panel of Experts Looking at Samples of Seeds (Beans Only) (Method F) 19
3.8. Validating Results of Individual Methods against DNA Fingerprinting 20
4. CONCLUSIONS AND IMPLICATIONS 24
4.1. Key Results and Lessons Learned 24
4.2. DNA Fingerprinting as a Method of Varietal Identification: Potential and Limitations ... 25
ANNEX 28
REFERENCES 35

LIST OF TABLES

TABLE **PAGE**

1. Methods of Varietal Identification Tested in the Pilot Countries 3

2. Characteristics of the Farmer and Number of Samples Collected in Zambia and Ghana to Test Different Methods 8

3. Classification of Farmer Samples and Accessions from Reference Library in to Unique Variety Clusters Based on DNA Fingerprinting: Results for Cassava in Ghana 10

4. Classification of Farmer Samples and Accessions from Reference Library in to Unique Variety Clusters Based on DNA Fingerprinting: Results for Beans in Zambia 12

5. Farmer Reported Names of Cassava Varieties Planted (Method A1): Results of the Cassava Varietal Identification Survey, Ghana 2013 13

6. Farmer Reported Names of Bean Varieties Planted (Method A1): Results of the Bean Varietal Adoption Survey, Zambia 2013 14

7. Farmer Reported Identification of Cassava and Bean Varieties by Type (Method A2): Results of the Cassava and Bean Varietal Identification Survey, Ghana and Zambia, 2013 .. 14

8. Matching Bean Varietal Names Elicited from Farmers Based on Showing Seed Samples (Method C) with Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013 16

9. Expert Reported Names of Cassava Varieties Based on Field Observations (Morphological Traits, Method D2a): Results of the Cassava Varietal Identification Survey, Ghana 2013 17

10. Expert Reported Groupings of Cassava Varieties by Type (Method D2b): Results of the Cassava Varietal Identification Survey, Ghana 2013 17

11. Expert Reported Groupings of Cassava Varieties by Type Based on Panel of Experts' Review of Photos Taken During Visit to Farmers' Field (Method E): Results of the Cassava Varietal Identification Survey, Ghana 2013 18

12. Matching Bean Varietal Names as Identified by Experts Based on Photographs of Seeds Taken by Enumerators during Farmer Surveys (Method E) with Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013 19

13. Expert Reported Groupings of Bean Varieties by Type Based on Review of Photos of Seeds (Method E): Results of the Bean Varietal Adoption Survey, Zambia 2013 19

14. Matching Bean Varietal Names as Identified by Experts Based on Seed Samples Collected by Enumerators during Farmer Surveys (Method F1) with the Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013 20

15. Expert Reported Groupings of Bean Varieties by Type Based on Panel of Experts' Review of Seed Samples (Method F2): Results of the Bean Varietal Adoption Survey, Zambia 2013 20

16. Measures of Effectiveness of Different Methods of Varietal Identification for Cassava in Ghana: Comparison of Outcomes against the Benchmark of DNA Fingerprinting Under the *Liberal* Assumption 21

17. Measures of Effectiveness of Different Methods of Varietal Identification for Cassava in Ghana: Comparison of Outcomes against the Benchmark of DNA Fingerprinting under the *Conservative* Assumption 22

18. Measures of Effectiveness of Different Methods of Varietal Identification for Beans in Zambia: Comparison of Outcomes against the Benchmark of DNA Fingerprinting 23

LIST OF ANNEX TABLES

TABLE **PAGE**

1. List of Cassava Released Varieties Included in the Reference Library for Ghana..... 29
2. List of Cassava Landraces Included in the Reference Library for Ghana 30
3. List of Bean Released Varieties Included in the Reference Library for Zambia..... 31
4. List of Farmer Reported Cassava Varieties Identified by them as Improved Varieties 32
5. List of Farmer Reported Bean Varieties Identified by them as Improved Varieties 33
6. List of Expert Reported Variety Names Identified by them as Improved Varieties..... 34

ACRONYMS

CGIAR	Consultative Group on International Agricultural Research
CIAT	International Center for Tropical Agriculture
DNA	Deoxyribonucleic acid
GBS	Genotype by Sequencing
GPS	Geographic Positioning System
IITA	International Institute of Tropical Agriculture
KASP	Kompetitive Allele Specific PCR
MSU	Michigan State University
PABRA	Pan African Bean Research Alliance
SNP	single nucleotide polymorphism
SPIA	Standing Panel on Impact Assessment of the CGIAR
ZARI	Zambia Agricultural Research Institute

1. INTRODUCTION

Since the pioneering research by Griliches (1958) on assessing the impact of hybrid corn adoption in the U.S. almost six decades ago, the interest in measuring the impacts of adoption of improved technology by farmers has expanded to include a gamut of agricultural technologies in both developed and developing country settings. Among the most widely assessed agricultural technologies in the developing country context is the adoption of improved varieties. These assessments have consistently reported that adoption of improved varieties and rapid varietal turnover increases productivity, income and other measures of welfare of farm households (Evenson and Gollin 2003; Zeng et al. 2015; Walker and Alwang 2015; Mathenge, Smale, and Olwande 2014).

Central to these assessments is the identification of improved varieties, estimating the adoption rate, and assessing varietal turnover. Most varietal adoption and impact assessment studies in the past have relied on either the low cost method of expert elicitation (e.g., Walker and Alwang 2015; Alene et al. 2009; Evenson and Gollin 2003) or the resource-intensive, but gold-standard method of conducting farm household surveys and eliciting this information directly from farmers (e.g., Zeng et al. 2015; Shiferaw et al. 2014; Kassie, Shiferaw, and Muricho 2011). However, despite their wide use, the reliability of these approaches has never been verified, leaving the bias and standard errors of these adoption estimates unknown.

Compared to the expert elicitation method, 'farmer elicitation' method can be fairly accurate in a setting where farmers are mostly planting seeds freshly purchased or acquired from the formal seed system as certified or truthfully labeled seed. In other words, the farm survey method can be effective if the seed system is well functioning and can effectively monitor the quality and genetic identity of varieties being sold by the seed vendors. However, in settings where the formal seed system is non-existent or ineffective, and farmers mostly rely on harvested grain (either from their own farms or acquired from other farmers or purchased from the market) as the main source of planting material, the reliability of estimating varietal adoption using this method is challenging (Yirga et al. 2015). By implication, it also makes the results of impact assessments based on those survey-based adoption estimates questionable.

The challenges stem from several confounding factors. These include biological factors such as the loss of genetic identity due to cross-pollination when seeds are recycled several seasons, and social factors such as: 1) farmers' inability to identify varieties by names; 2) the inconsistency in the names of the varieties as identified by the farmers and what is in the variety registration list (i.e., varieties may have locally adapted names); and 3) farmers' lack of understanding of what is an *improved/modern variety* vs. *unimproved/traditional* variety or inability to distinguish between different types of hybrids and varieties.

To tackle these potential problems in collecting variety-specific adoption data from farmer-level surveys, requires more data collection either by including follow-up questions in the survey instrument on varietal traits and/or collecting sample materials (i.e., photos, samples of seeds) from the farmers to be later verified by experts. There are also other non-household based adoption tracking methods that may be feasible for some crops such as through record keeping and collection of minimum data from major actors along the seed supply chain or Deoxyribonucleic acid (DNA) fingerprinting of grain samples collected in major markets

after the harvest.[1] Each of these potential approaches to track varietal adoption has implications on the cost, logistics, and timing of conducting the survey. However, despite the costs, the accuracy and credibility of these different methods to identify a variety to be able to estimate its adoption still remains an open question. The objective of this paper is to precisely address this unanswered question. It reports the results of two pilot studies conducted in Ghana and Zambia to test different approaches of collecting variety-specific adoption data and to validate them against the benchmark of DNA-fingerprinting to determine which method is most accurate in identifying varieties grown by farmers.

DNA fingerprinting, which is increasingly used by plant breeders, offers a reliable method to accurately identify varieties grown by farmers and thus serves as a benchmark against which traditional or other innovative methods of varietal identification can be evaluated. However, despite this advantage, the use of DNA fingerprinting as part of adoption surveys is non-existent or limited to few recent attempts, mostly at pilot scales (Rabbi et al. 2015; Kosmowski et al. 2016; Floro et al. 2016). With the cost of genotyping expected to decline rapidly, the use of DNA fingerprinting may be pursued as the main method of varietal identification to track and monitor the adoption of improved varieties. Thus, a secondary objective of this paper is to derive lessons and implications for scaling up the use of DNA fingerprinting method for varietal identification.

We first describe the different methods of varietal identification evaluated in the two countries for two different crops. We also describe the sampling approach used, the logistics of collecting samples, and the cost of doing DNA fingerprinting in these two pilot studies. Results of varietal identification using different methods, including DNA fingerprinting, are presented next, followed by conclusions and implications for future adoption and impact studies.

[1] Note that these other methods may have limited applications because of the problem of external validity (or sample 'representativeness'). For example, samples collected in markets may not be representative of total production as some varieties may be adopted specifically for sale and some for home consumption, and sales are often correlated with asset endowments which may also affect the propensity to adopt improved varieties.

2. METHOD AND APPROACH

2.1. Methods of Varietal Identification Evaluated

Two pilot studies were conducted—one in Ghana for cassava (*Manihot esculenta*) and the other in Zambia for beans (*Phaseolus vulgaris*)—to test the effectiveness of alternate methods of tracking varietal adoption using farm household surveys. Both of these pilots involved a multi-disciplinary team of experts (i.e., breeders, geneticists, social scientists, and economists) from national agricultural research systems, international agricultural research institutes, and U.S. universities.

Six methods of tracking varietal adoption using farm household surveys were evaluated across the two crop-country combinations (Table 1). These methods can be grouped into two types—farmer elicitation methods (methods A-C) and expert elicitation methods (methods D-F) (Table 1). The protocol followed for each of these methods is as follow:

Method A: As part of the survey instrument, farmers were asked to provide the name(s) and type (improved vs. local) of varieties planted in the current planting season (for cassava) or the last completed season (for beans). Thus, we were able to analyze the information in two ways: farmers' reported name of the variety grown (Method A1) and farmers' classification of the variety grown as improved or local (Method A2). In both the cases additional information on the source of the first and current planting materials and number of years a variety was grown by the farmer was collected.

Method B (tested only for cassava): As part of the survey interview, farmers were asked about specific varietal characteristics by showing a series of photographs. In the case of cassava, pictures of cassava plants depicting 11 different morphological characteristics were shown and farmers were asked to identify the characteristic (e.g., color, size, shape) that best matched the characteristics of the variety he/she was growing. Farmer responses were later matched with the variety specific characteristics of all the accessions included in the reference library (as catalogued by the cassava breeder from Ghana and the International Institute of Tropical Agriculture (IITA) to identify the variety cultivated by the farmer.

Table 1. Methods of Varietal Identification Tested in the Pilot Countries

	Methods	Ghana (cassava)	Zambia (beans)
V	DNA fingerprinting (used as a benchmark to compare/validate other methods)	X	X
A	Farmer elicitation by: 1) name and 2) type of variety) \a	X	X
B	Farmer elicitation based on a series of photographs of plants	X	
C	Farmer response on type of variety he/she had planted that match seed samples presented by the enumerators		X
D	Trained enumerators/experts visiting the field and: 1. Recording observations on varietal characteristics (phenotyping); and 2. Identifying the variety by name (D2a) and type (D2b) based on observation (phenotyping)	X	
E	Taking photos of the plant in the field or seeds harvested by farmers for latter identification by experts (i.e., breeders, agronomists, etc.) by name (E1) and type (E2)	X	X
F	Collecting harvested seeds from farmers for latter identification by experts (i.e., breeders, agronomists) by name (F1) and type (F2)		X

Source: Authors for all tables.

Method C (tested only for beans): This method involved showing the farmer seed samples representing different varieties and asking him/her to identify the seed sample that matched the varieties grown on their farm. To implement this method, all the released bean varieties were organized in small pockets of plastic bags that were easy to carry by enumerators. As part of one of the survey modules, farmers were asked whether the variety he/she is growing matched any of the variety in this sample. If the response was yes, the enumerator noted down the sample code and latter matched with the name of the variety corresponding to that sample. If the farmer's response was no, the variety grown by the farmer was interpreted as *other traditional/local landraces*.

Method D (tested only for cassava): A trained cassava expert (i.e., field technician from the cassava breeding program) was included as part of the survey team and visited the cassava fields enumerated in the household survey to test two closely related methods of varietal identification. The first method (D1) involved the expert recording his/her observations on 11 morphological characteristics (same ones included in Method B) of a representative plant corresponding to each variety planted on that field as identified by the farmer. This information recorded by the expert was later matched with the variety specific characteristics of all the accessions included in the reference library (as catalogued by the cassava breeder) to identify the variety cultivated by the farmer. A second variation of this method (D2) was varietal identification (by name (D2a) and type (D2b)) recorded by the visiting expert himself or herself based on observation of the plant in the field.

Method E: This method consisted of taking photographs of the plant in the field (in the case of cassava) or a sample of harvested seeds (in the case of beans) and later using these pictures for varietal identification by a panel of experts, by name (E1) and type (E2).

Method F (tested only for beans): Consisted of collecting seed samples of varieties grown by the farmer for later identification by breeders or other bean experts, by name (F1) and type (F2). In the case of beans, this method is an extension of the step involved in doing the DNA fingerprinting analysis for varietal identification. The seed samples collected for that purpose were used to seek expert elicitation on varietal identification.

2.2. DNA Fingerprinting

In both the pilots, DNA fingerprinting (method V) was used as a validation method against which alternative approaches were evaluated. This involved first establishing a reference library of DNA fingerprints, and then collecting samples (plant tissues or seeds) during the household/farm surveys and genotyping them using the same or a sub-set of markers used to establish the reference library. In the case of cassava, a total of 64 accessions of released varieties (n=18) and popular landraces (n=46) were included in the reference library (see Annex 1 and Annex 2). Samples of these accessions along with the samples collected from farm surveys were all genotyped at 56,849 single nucleotide polymorphisms (SNP) loci. Genetically identical sets of clones were then identified by using distance-based hierarchical clustering and model-based maximum likelihood admixture analysis (Rabbi et al. 2015).
In the case of beans, 13 accessions specific to Zambia (including 11 released varieties listed in Annex 3 and two landrace Kablanketi market classes), ~25 farmer collected samples, and 698 accessions from the East/Southern Africa region (that were genotyped as part of another project by the International Center for Tropical Agriculture (CIAT)) were included in the reference library as the background materials to compare the samples collected during the household surveys. The farmer samples were genotyped using 66 Kompetitive Allele Specific PCR (KASP) assays/markers selected as a sub-set of 776 SNP KASP markers used for the

reference library. The 66 SNP assays were made up of 4 groups, each of which has more or less the same power to differentiate the 11 released varieties from each other and from the background genotypes (Raatz 2015).

2.3. Sampling and Data Collection

2.3.1. Ghana

In Ghana, the pilot study was conducted in three regions that account for 61% of cassava production in the country in 2010 (Angelucci 2013; FAOSTAT 2014). The three study regions included Brong Ahafo, Ashanti, and Eastern. A total of 500 households were targeted for the survey using a multistage cluster sampling method. Only districts with more than 5,000 ha under cassava cultivation were identified in the sampling frame. In the first stage, five-eight districts per region were randomly selected from this sampling frame. In the second stage, five villages per district were randomly selected. Finally, five cassava farmers from each village were selected based on a random start and then skipping x = N/5 (where N = approximate number of households in the village) number of houses until the target number of cassava farmers were reached.

A total of 495 cassava growing households distributed across 100 villages from 20 districts in the three study regions of Ghana were surveyed in October-November 2013. The survey was coordinated by a research team led by the cassava breeder from the Council for Science and Industrial Research-Crops Research Institute of Ghana and a socio-economist from the Agriculture Innovation Consult. The survey team consisted of enumerators who were in-charge of completing the household modules, a cassava expert in-charge of completing the field survey module, and a DNA sampling expert in-charge of collecting, labeling and storing the plant tissue samples as per the protocol established with the help of researchers from IITA. In each household, the field with the most number of cassava varieties grown by the farmer were visited to collect the leaf samples for DNA analysis and to collect information/ pictures required to test methods D and E described above. The GPS coordinates of the field were recorded and farmers were asked to identify plants representing each of the varieties they reported growing in the current season during the household survey interview. During the field visits if the cassava expert found natural variation in the observed characteristics of a variety, they were instructed to collect leaf samples from each variation of a variety observed (and label them as variations of variety x). Apical leaf samples were collected from one plant representing each variety (and its variation) and preserved in silica gel for transportation to a central laboratory at IITA in Ibadan, Nigeria for DNA extraction. The extracted DNA samples were then shipped to Cornell University's Genomic Diversity Facility at the Institute of Biotechnology for analysis using the genotyping-by-sequencing (GBS) method. Data interpretation and analysis for varietal identification based on the GBS results was led by researchers at IITA.

2.3.2. Zambia

In Zambia, the pilot study was designed to take advantage of an already planned bean varietal adoption and impact study by the Zambia Agricultural Research Institute (ZARI) with support from Pan African Bean Research Alliance (PABRA) and CIAT (Hamazakaza et al. 2014). The study was conducted in Muchinga and Northern Provinces of Zambia. These provinces were purposively selected because of their importance in bean production, accounting for about 70% of the area under beans in Zambia, and because most of the prior seed dissemination efforts were concentrated in this part of the country.

A total of seven districts were purposively selected (again based on the importance of the bean crop) from the two provinces: four districts (Kasama, Mbala, Mporokoso, Mpulungu) in the Northern Province and three (Chinsali, Mpika, Nakonde) in the Muchinga Province, which together represent 59% of the total bean area in Zambia. After the districts were selected, a two-stage cluster sample selection method was used. In the first stage, villages were randomly selected from each district according to the proportion of villages within the selected districts in each province. In the second stage, six households were systematically selected within each village. A sample of 400 farmers across 67 villages was determined based on the available budget. Thus, 41 and 26 villages were selected in the Northern and Muchinga Provinces, respectively and 6 farmers per village were surveyed to get a total sample size of 402 farm households. To select the households, a systematic random sampling procedure was followed. The village register list obtained from the local headman or village secretary served as the sampling frame and each household in this list was numbered sequentially. The first household was selected at random from this list, and the remaining five households were chosen at fixed interval $x = N/6$ (where N = number of households on the village list) until the target number of bean farmers was reached.

The survey was implemented between August-September 2013 and the information collected refers to the 2012-2013 agricultural season (December 2012-April 2013). The enumerators were trained by a research team from ZARI, MSU, and CIAT on how to use the instruments for household- and village-level data collection, how to take photographs of the seed samples (method E), and how to implement the protocol for collecting 10-15 seeds of each variety the farmer had harvested in the 2012-2013 agricultural season and labeling them for proper tracking (methods F and V). Each enumerator received a set of seed samples representing ten different released varieties that was presented to the farmer to facilitate in variety identification (method C). Each small plastic bag containing these seeds had a code and only the supervisors knew which code belonged to which variety.

For DNA fingerprinting (method V), the seed samples collected from the farmers were germinated by the ZARI bean breeder at the Kasama research station in May-June 2014 (after these were used for implementing method F). With the help of a CIAT technician, leaf tissue samples from young germinated bean plants were collected in a 96 well-plate leaf sampling kits and shipped to LGC Genomics lab in U.K. for genotyping. For quality control and sample verification purpose, all the seed samples were shipped to Michigan, and a sub-set of the seed samples were germinated at Michigan State University with the help of the bean breeding program. Same procedure was used to collect the leaf samples in a 96 well-plate sampling kits and shipped to LGC Genomics lab for genotyping. All the farmer samples were genotyped using 66 SNP markers (or KASP assays) that were identified specifically for this study (as described above).

In both the countries, household level questionnaires collected information on household characteristics, farm characteristics, varietal identification questions corresponding to methods A-C, and variety-level questions on preferences, use, like/dislike characteristics, etc. In both the countries, a community level questionnaire was also completed to collect some community level characteristics that can explain varietal adoption outcomes. In the case of Ghana the household level data were collected using *SurveyBe*, a computer assisted personal interviewing (CAPI) method. In Zambia, the survey was conducted using the paper-based personal interviewing method. Farmers' participation in both the surveys was voluntary and they were fully informed on the survey objectives and how they were selected to participate in the survey.

2.4. Varietal Identification Using Expert Elicitation Based on Photographs and Seed Samples

After the survey, the seed samples and photographs collected in Zambia, and photographs of plants taken in Ghana were used in varietal identification (methods D-F) by a panel of crop experts who were familiar with the varieties grown in the study area. For beans, the breeders and extension staff from the study districts were invited to ZARI's Misamfu Research Station in Kasama to participate in the identification of varieties using seed samples (day 1) and photos (day 2). This expert panel meeting in Zambia was conducted in March 2014. For cassava, the expert elicitation panel discussion was organized at CRI in August 2014 and included cassava breeders, experts and field technicians from CSRI-CRI, IITA, and the University of Cape Coast. In both the cases, the overall elicitation process was facilitated by a socio-economist either from IITA (in the case of cassava) or from MSU (in the case of beans) who did not participate in varietal identification.

There were some differences in the implementation of expert elicitation method in the two countries. In the case of Zambia, the bean samples were identified by names and if the experts could not name a variety, *no name* was recorded as their answer. On the other hand, in the case of cassava, variety was mostly identified as improved, not improved, or unknown. In addition, in the case of cassava the opinion on the type of variety based on the photo identification method was recorded separately by each of the seven experts. However, in the case of beans, after some exchange of opinions a consensus about the name of the variety was reached and only one name corresponding to a sample was recorded. If the experts could not identify a variety by name, no name was recorded as their answer. In some instances, when a variety was identified as a *mixture* (which was common in Zambia), each of the varieties within this mixture were evaluated separately (i.e., each was given a name).

Next, we present descriptive statistics and make a comparison of the adoption estimates from the different methods evaluated in this study. As shown by the results presented below, there are large differences between the estimates of adoption rates obtained by these different methods, compared to fingerprinting results. Some of these methods were very complex to operationalize and fewer varieties could be identified with these methods. Moreover, it was more challenging to identify cassava varieties than bean varieties.

It should be noted that the term *adoption rate* as used in this paper refers to data points as a proportion of total number of observations of farmer reported varieties. It is not a measure of adoption of varieties in the classical sense of adoption as a percentage of farmers or total area planted to a given crop. For the purpose of this study, which is to test the effectiveness of different methods, the unit of comparison is a data point, rather than a farmer or a hectare of land. The adoption estimates refer to percentage of farmer samples that match *released varieties* included in the reference library. The terms *local variety* or landraces are used interchangeably and refer to varieties that have not been subjected to any rigorous selection or breeding efforts by a research program.

3. RESULTS

Table 2 presents key descriptive statistics of the 402 bean-growing households surveyed in Zambia and 495 cassava-growing households in Ghana. On average a sampled household in Zambia cultivated beans on fewer plots than cassava growers in Ghana (1.4 plots vs. 2 plots, respectively).[2] A typical farmer in the study area of Zambia had cultivated two varieties of beans in the 2012-2013 season. This resulted in a total of 863 bean varietal observations (or data points) for which different methods of varietal identification could be applied. Out of these, 855 were genotyped. In the case of cassava, the average number of cassava varieties was reported to be 1.9 per household, resulting in 924 total number of cassava varietal observations across 495 households. Out of this total number of varieties documented, 914 were genotyped.

In the case of beans, since the survey was conducted several months after the harvest season, many farmers had no seed stock left to share with the enumerators. This is the reason why the number of samples genotyped for beans was less than the total number of varieties documented in the household survey. Moreover, due to technical glitches and quality issues, the database of photographs available for varietal identification for both beans and cassava was substantially lower than the number of seed/plant samples collected (Table 2).

3.1. Classification of Reference Library Materials and Farmer Samples into Variety Groups Based on DNA Fingerprinting

Tables 3 and 4 present the results of DNA fingerprinting (method V) and its implication on the classification of farmer samples and library samples into unique variety groups for cassava and beans. As reported in Table 3, the DNA analysis identified eleven unique groups of cassava varieties and twelve groups of hybrids or admixtures.

Table 2. Characteristics of the Farmer and Number of Samples Collected in Zambia and Ghana to Test Different Methods

Details	Zambia (Beans)	Ghana (Cassava)
Number of farmers surveyed	402	495
Average number of plots on which the crop was planted (range)	1.36 (1-4)	2.05 (1-25)
Average number of varieties planted per household (range)	2.08 (1-6)	1.92 (1-7)
Average number of varieties planted per plot visited (range)	--	1.85 (1-5)
Number of varietal data points reported in farmer and market survey (for method A and B)	863	924
Number of samples genotyped (DNA fingerprinted) (for method V)	855	914
Number of samples photographed and available for varietal identification (for method E, F)	736	792

[22] The average bean plot size of farmers sampled in Zambia was 0.66 ha. For Ghana we only collected information on the field size where most number of cassava varieties were planted. The average size of such fields across our sample was 1.3 ha.

Most of the samples from farmers' fields were classified into variety groups 1 and 2 and multi-ancestry clones of group 1. From the 11 variety groups, eight could be easily classified as local or released varieties and three (i.e., variety groups 2, 3, 4) could not, as they contained genetic materials from library accessions that fall under both local and released varieties. For variety group 4, cassava breeders were confident that Afisiafi was an output of the breeding program, and thus this group is classified as released variety, despite the three accessions (i.e., ABUSUA, MONICA, and UCC2001_449) that fall in this group were considered to be landrace materials. For the other two groups (i.e., groups 2 and 3), varieties IFAD, UCC, and Nkobam were improved landraces released in Ghana and share genetically identical fingerprints with several local accessions included in the library. There were no strong convictions on whether to consider farmer collected samples that fall in this group as released varieties (representing some improvement by the breeding program) or local landraces. Given this ambiguity, for the comparative analysis we present two scenarios: a *conservative scenario*, where farmer collected samples that matched variety groups 2 or 3 are classified as local varieties, and a *liberal scenario*, where they are classified as released/ improved varieties. Finally, farmer varieties that did not match any genetic material from the reference library were classified as *unknown varieties* (or varieties not known to be released varieties). Under the liberal scenario 284 (out of 914) data points corresponding to farmer collected samples are classified as released varieties, whereas under the conservative scenario only 40 (out of 914) data points are classified as released varieties (Table 3).

In contrast, for beans, the identification of groups was much simpler (Table 4). A total of 12 variety groups were identified, and a group called *other* was also identified when there was no match between the samples collected from farmers with the reference library. This other group is classified as local varieties. From the 12 variety groups that were identified, 11 were classified as released (i.e., improved) varieties and one as local variety.

Most of the samples from farmers were classified into variety groups other (690 out of 855 data points), group 3 (124 data points) and group 12 (24 data points). Overall, 141 out of total 855 data points of farmer and market bean samples match the DNA fingerprints of released varieties across groups 3 (Kabulangeti set 3), 7 (Lukupa), 8 (Lwangeni), and 9 (Lyambai). For seven of the released varieties included in the reference library (i.e., Chambeshi, Kabale, Kalambo, Kalungu, Kapisha, Mbereshi, and Sadzu) there was no match with the samples collected from the farmers. Of the 690 samples marked other, many samples look identical to the Fungula cluster included in the background materials (i.e., Fungula_10008, Pomba_10003, and Tamba 10016). This appears to be the most popular genotype, after the Kabulangetis. Several farmer samples were found identical to other Zambian background genotypes included in the reference library such as ZM 4465_9015, Lusaka_10018, ZM 3616_9002, and ZM 7295_9030. These appear to represent popular but diverse sets of landraces in Zambia.

The effectiveness of varietal identification using alternative methods A to F described Table 1 is evaluated against this benchmark of varietal groups identified based on DNA fingerprinting results. We first present the results of how varieties grown by farmers are identified using each of these alternative methods.

Table 3. Classification of Farmer Samples and Accessions from Reference Library in to Unique Variety Clusters Based on DNA Fingerprinting: Results for Cassava in Ghana

Unique variety group based on DNA finger-printing	# of accessions from farmer samples classified under a variety group	Accessions from reference library that fall in the variety group		Classification of farmer samples under two assumptions	
		Released varieties	Landraces/Local varieties	Liberal	Conser-vative
Variety 1	208	No match	(12) ADW2000_003; ADW2000_004; ANKRA; BOSOMENSIA_1; DEBOR 1; DEBOR_KAAN; DMA2000_002; DMA2000_66; KSI2000_126; OFF_2000_019; OFF_2000_023; UCC2000_111	Local variety	Local variety
Variety 2	158	(2) IFAD; UCC (improved landraces released in Ghana)	(7) TUMTUM; DWA2000_070; ELISHA; WCH2000_020; KW_2000_010; KWANWOMA; OFF_2000_134	Released variety	Local variety
Variety 3	65	(1) NKABOM *(improved landrace released in Ghana)*	(1) DEBOR 2	Released variety	Local variety
Variety 4	17	(1) AFISIAFI	(3) ABUSUA; MONICA; UCC2001_449	Released variety	Released variety\a
Variety 5	58	No match	(2) ADE2000_182; DMA2000_031	Local variety	Local variety
Variety 6	36	No match	(2) KW_2000_148; UCC2001_399	Local variety	Local variety
Variety 7	20	No match	No match	Unknown variety	Unknown variety
Variety 8	21	No match	(2) BANKYE_BRONI UCC20001_464;	Local variety	Local variety
Variety 9	13	No match	No match	Unknown variety	Unknown variety
Variety 10	33	No match	No match	Unknown variety	Unknown variety
Variety 11	10	No match	No match	Unknown variety	Unknown variety
Hybrids (or Admixtures)					
50% ancestry from variety 1	17	No match	No match	Unknown variety	Unknown variety
50% ancestry from variety 2	11	No match	No match	Unknown variety	Unknown variety
50% ancestry from variety 3	19	(3) ESSAM_BANKYE; BANKYE_HEMAA ; TEKBANKYE; DOKU_DUADE	No match	Released variety	Released variety

Unique variety group based on DNA finger-printing	# of accessions from farmer samples classified under a variety group	Accessions from reference library that fall in the variety group		Classification of farmer samples under two assumptions	
		Released varieties	Landraces/Local varieties	Liberal	Conser-vative
50% ancestry from variety 4, group 1	8	No match	(2) BRONI; KW2000_181	Local variety	
50% ancestry from variety 4, group 2	2	(3) NYERIKOGBA; ABASA_FITAA; OTUHIA	No match	Released variety	Released variety
50% ancestry from variety 5	12	No match	No match	Unknown variety	Unknown variety
50% ancestry from variety 6	34	No match	No match	Unknown variety	Unknown variety
50% ancestry from variety 8	21	No match	(4) 12_0236; 12_02Y5; CONGO_BATIALION; ESIABAYAA	Local variety	Local variety
50% ancestry from variety 9	29	No match	No match	Unknown variety	Unknown variety
50% ancestry from variety 11	5	No match	(2) KW_2000_030; UCC2001_249	Local variety	Local variety
Multi-ancestry clones, group 1	115	No match	(11) 12_0197; ADW2001_051; AFS_2000_050; ANKRA_10_003; AW3_10_008; AW3_10_011; BOSOMENSIA_2; CONGO_BATIALION; DEBOR_BEPOSO; OFF_2000_037 WCH2000_011	Local variety	Local variety
Multi-ancestry clones, group 2	2	(6) BRONI_BANKYE; AMPONG; FILINDIAKONIA; BANKYE_BOTAN; SIKABANKYE; AGBELIFIA	No match	Released variety	Released variety
Total	914	18	46		

\a As noted in the Report, Afisiafi is considered a distinct output of the breeding program, and thus this group is classified as a released variety under both the scenarios.

Table 4. Classification of Farmer Samples and Accessions from Reference Library in to Unique Variety Clusters Based on DNA Fingerprinting: Results for Beans in Zambia

Unique variety group based on DNA fingerprinting	# of accessions from farmer samples classified under this variety group	Accessions from reference library that fall in this variety group		Classification of farmer samples that fall in this cluster group
		Released varieties	Landraces/ Local varieties	
Variety 1	0	Chambeshi		Released variety
Variety 2	0	Kabale		Released variety
Variety 3	124	Kabulangeti, set 3		Released variety
Variety 4	0	Kalambo		Released variety
Variety 5	0	Kalungu		Released variety
Variety 6	0	Kapisha		Released variety
Variety 7	13	Lukupa		Released variety
Variety 8	3	Lwangeni		Released variety
Variety 9	1	Lyambai		Released variety
Variety 10	0	Mbereshi		Released variety
Variety 11	0	Sadzu		Released variety
Variety 12	24		Kablanketi	Local variety
Other	690	No match	No match	Unknown variety

3.2. Classification of Varieties by Farmers' Reported Name and Type of Variety (Method A)

Tables 5-7 presents the results of varietal identification based on method A which includes farmers' reported name of the variety grown (method A1) and farmers' classification of the variety grown as improved or local (method A2). As the results indicate, the variability in the names of varieties reported by farmers is large, particularly for cassava since farmers reported a total of 180 variety names compared to approximately 57 names reported by bean farmers in Zambia. Furthermore, 20 cassava samples collected from farmers' fields corresponded to 8 names of released cassava varieties (Table 5) and 34 bean samples corresponded to 5 names of released bean varieties (Table 6). The most commonly reported improved cassava variety was Bankye Hemaa and the most commonly reported bean variety was Kabulangeti. The estimates of adoption (as measured by data points classified as released varieties as a proportion of all the data points) using this method (i.e., method A1) is 2.2% for cassava and 3.97% for beans.

Farmers were also asked to classify the varieties they grew as improved or traditional varieties (method A2), based on their knowledge and belief system. As reported in Table 7, the estimates of adoption rates were considerably higher when using this method compared to using the names of the varieties. Using this method, for cassava, 5.6% of the samples were classified as improved/released varieties and for beans, this estimate was 13.2%. Clearly, many farmers were reporting names of local varieties but considering them as improved varieties both for cassava (Annex 4) and beans (Annex 5). In the literature, most adoption studies have relied on this method (A2) or a combination of method A1 and A2 to estimate adoption rates and for defining the varietal technology adoption variable that is often used as a dependent variable or an explanatory variable in econometric analyses.

Table 5. Farmer Reported Names of Cassava Varieties Planted (Method A1): Results of the Cassava Varietal Identification Survey, Ghana 2013

Group based on names	Variety names as reported by farmers	Frequency	Total
Released/Improved varieties	BANKYE_HEMAA	7	20
	NKABOM	4	
	AFISIAFI	3	
	ABASA_FITAA	2	
	AMPONG	1	
	ESAM_BANKYE	1	
	NKAKOM_BANKYE	1	
	SIKA_BANKYE	1	
Not matching names of any released varieties	DEBOR	88	894
	ANKRA	86	
	BANKYE_KOKOO	49	
	BOSOMENSIA	49	
	ABENWOHA	39	
	ESIABAYAA	32	
	KOTEE	28	
	BANKYE_FITAA	23	
	TUAKA	21	
	AGRIC	19	
	AMPENKYENE	17	
	AFIA_KOFIE	13	
	BANKYE_TUMTUM	13	
	BANKYE_BOADIE	11	
	KENTENMA	12	
	ESIABAAYEWA	10	
	MEDUMEKU	10	
	Other varieties with frequency 5-9	127	
	Other 138 varieties with frequency <5	228	
	UNKNOWN	19	
Total	180 variety names		914

Table 6. Farmer Reported Names of Bean Varieties Planted (Method A1): Results of the Bean Varietal Adoption Survey, Zambia 2013

Group based on names	Variety name reported by farmers \a	Frequency	Total
Released/ Improved varieties	Lwangeni	1	34
	Lyambai	1	
	Kapisha	1	
	Kabulangeti \b	30	
	Lukupa	1	
Not matching any released varieties	Local Kabulengeti \b	195	821
	Pomba	38	
	Lusaka	64	
	Solwezi	25	
	Market class – red	21	
	Market class – yellow	43	
	Market class – white	195	
	Market class - yellow and white	101	
	Market class - red & white	1	
	Market class – black	1	
	Market class – brown	3	
	Mixed	77	
	Other names (40 unique names)	55	
	No name	2	
All	~57 unique names		855

\a Several varieties were named by the seed color such as *red yellow, white*, etc. These are akin to bean market classes.
\b The grouping of farmer reported kabulengeti into released variety and local is based on farmer response to a follow-up question on whether a variety was improved or local.

Table 7. Farmer Reported Identification of Cassava and Bean Varieties by Type (Method A2): Results of the Cassava and Bean Varietal Identification Survey, Ghana and Zambia, 2013

	Frequency	
Variety type reported by farmers	Cassava	Beans
Improved	51	106
Local	796	612
Don't know/no response	67	84
Total	914	802

3.3. Classification of Varieties by Farmer Reported Characteristics of the Plants (Cassava Only) (Method B)

As explained above, for this method farmers in the Ghana study were asked about specific characteristics of each variety grown by showing a series of photographs. The 11 morphological characteristics shown to farmers correspond to different stages of plant growth and included questions about: the color of the apical leaves and pubescence of apical leaves three months after planting; the petiole color, leaf color and shape of central leaflet six

months after planting; and the color of stem, growth habit of stem, root shape, root color for outer skin, root color for inner skin, and color of root pulp nine months after planting.For each of these characteristics there were specific codes used to record the answers (e.g., for color of apical leaves: 3=light green, 5=dark green, 7=purple green, 9=purple) and these codes corresponded to some scoring system normally used by breeders to characterize cassava varieties. The same method was used to record the codes for the reference library materials across the same 11 morphological traits. Responses to the questions related to the morphological traits were then used to create an 11-digit number where each digit represented the score for a trait in the sequence in which these codes were recorded. The idea behind this method was that the released varieties would be different on at least one of these morphological traits, and, thus, the 11-digit number generated for each released variety would be unique (like a fingerprint), and, therefore, could be used as a benchmark to identify farmer varieties based on data points for which this 11-digit code exactly matched with a variety included in the reference library.

Among all the methods of varietal identification tested in this study, this was the most time and resource intensive and the least effective method. Both breeders and farmers had difficulty providing information on all the characteristics needed to generate the 11-digit code. For the cases where this information was available and the code was generated, only two data points matched the codes in the library samples—one each for varieties UCC and Bankye Broni.

Thus, although varietal identification of cassava based on visual characteristics of the plant has long been used by breeders, and such information is published for released varieties to help farmers identify improved varieties, in practice, this approach was difficult to operationalize. We thus consider the potential for scaling up this method for varietal identification as part representative household surveys to be limited.

3.4. Classification of Varieties by Farmers' Based on Matching their Variety with Seed Samples Shown (Beans Only) (Method C)

As one of the low-cost and operationally easy methods of varietal identification, farmers were shown samples of seeds of all released varieties in small plastic bags, coded with numbers. They were then asked to point to the sample that matched the variety they grew during the season of reference and their response was recorded. If a variety did not match a seed sample that the enumerator had shown, it was assumed to be local.

Based on this method, a total of 553 observations, (or 71%) of farmer samples were identified with one of the samples of released varieties shown to farmers (Table 8). This estimate of use of improved varieties (71%) based on this method is substantially higher than the estimates derived from any other methods tested, including DNA finger printing. Varieties Kabulangeti and Kulungu were the most commonly matched bean varieties by the farmers based on this method (Table 8). However, farmers did not distinguish between the improved and local Kabulangeti and in some cases, they reported different names for the same variety. In 15 cases, the variety grown matched more than one sample of IVs shown.

Table 8. Matching Bean Varietal Names Elicited from Farmers Based on Showing Seed Samples (Method C) with Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013

Variety name as identified based on farmer matching planted varieties with seed samples	Variety name in the reference library included for DNA fingerprinting (benchmark)					Does not match any genotype	Total
	Lwangeni	Lyambai	Lukupa	Kabulengeti (improved)	Kabulengeti (local)		
Lwangeni	1	0	1	0	0	33	35
Lyambai	0	0	0	0	0	4	4
Lukupa	0	0	5	0	0	4	9
Kabulangeti (improved)	0	0	1	57	15	134	207
Kabale	0	0	0	2	0	10	12
Kalungu	0	0	0	25	5	177	207
Kapisha	0	0	0	1	0	19	20
Kalambo	0	0	1	2	0	36	39
Chambeshi	0	0	1	0	0	2	3
Mbereshi	0	0	0	0	0	2	2
Matches multiple	1	0	0	1	2	11	15
Does not match any released variety	1	1	3	18	3	204	230
Total	3	1	12	106	25	636	783

3.5. Classification of Varieties by Cassava Experts Visiting the Field and Recording Observed Morphological Traits (Cassava Only) (Method D)

For this method, cassava experts visited the field and collected information on observable morphological characteristics (same characteristics listed for method B), which were later used to generate a unique code to match with the codes of library materials (method D1; similar to method B). Further, experts were asked to use that information in situ to identify the observed variety by name (method D2a) and type (method D2b).

As with method B, it was difficult to record all the plant characteristics corresponding to different stages of plant growth during this single visit. For the cases where this information was recorded to generate the 11-digit code, only two matched the library: one each for varieties DMA2000-066 and 12-0236, both landraces. Thus, using this method, the adoption rate of IVs was zero percent.

Cassava experts visiting the field reported 86 names of varieties based on the morphological characteristics they observed during the visit (method D2a). From these, only 33 names matched the name of a variety released in Ghana (Table 9), suggesting a 3.6% adoption rate of IVs. Implementing this method proved difficult as for 299 observations, experts could not provide a name for the variety, and there were 39 observations with missing data.

Although identification of cassava varieties by name using this method was not possible for all the observations, experts who visited the cassava fields were able to identify the variety by type for most varieties they observed, except for 62 cases (Table 10).

Table 9. Expert Reported Names of Cassava Varieties Based on Field Observations (Morphological Traits, Method D2a): Results of the Cassava Varietal Identification Survey, Ghana 2013

Group based on names	Variety name as given by the expert\a	Frequency
n.a.	No name	299
Landrace	Debor	98
Landrace	Bosomensia	62
n.m.	Bankye Kokoo	43
n.m.	Aben Woha	36
n.m.	Bankye Pole	36
Landrace	Ankra	23
n.m.	Afia Kofie	19
Landrace	Esiabayam	19
Landrace	Esiabayewaa	18
n.m.	Ampenkyene	17
Improved	**Bankye Hemaa**	13
Improved	**Afisiafi**	10
Improved	**Nkabom Bankye**	10
n.m.	Tuaka	10
n.m.	Akosua Tuntum	8
n.m.	Kentenma	8
n.m.	Ahanta	6
n.m.	Tuabodom	6
n.m.	Bankye Seth	5
n.m.	Biambashie	5
n.m.	Other 66 varieties	124
n.m.	Missing data	39
Total		914

\a Variety names that match the released varieties in Ghana are bolded.
n.a.= not applicable; n.m.= no match.

Table 10. Expert Reported Groupings of Cassava Varieties by Type (Method D2b): Results of the Cassava Varietal Identification Survey, Ghana 2013

Variety type reported by field visiting expert	Frequency
Improved	47
Local	805
Don't know/missing information	62
Total	914

Based on this method (method D2a) a total of 47 observations were classified by experts as improved varieties, suggesting an adoption rate of 5.1%. However, when the names associated with varieties classified as IVs were compared with the varietal identity based on DNA results, only six named varieties matched released varieties (Annex 6), suggesting a low rate of accuracy of varietal identification based on this field based observation of varietal characteristics.

3.6. Classification of Varieties by Experts Looking at Pictures of Cassava Plants and Pictures of Bean Seeds (Method E)

Enumerators who visited the cassava field also took pictures of the plants associated with different morphological characteristics (e.g., leaf and branch colors, branching habits, root color, etc.), and these were later used by a panel of cassava experts (breeders and agronomists) to provide a name of the variety and type (method E). Experts were not able to agree on names for the varieties using this method. Experts expressed concerns about picture quality and the fact that the pictures were all for one-time shots of matured plants and did not capture earlier stages of vegetative growth. Moreover, even though the breeders agreed that four to five pictures should have been taken for each plant, some experts later said that more information was required for them to identify the varieties by name. Despite these limitations, they were able to give their opinion on whether they thought the varieties were improved or local. As can be seen in Table 11, the panel of experts identified 142 data points as improved, suggesting that 15.5% of cassava samples collected from farmers' fields were improved varieties according to this method of varietal identification.

Similarly, for beans, enumerators took pictures of the seed samples that were grown by farmers and a panel of experts used these to identify the seed sample by a name and type of bean variety. Given that not all farmers had seed samples to show, this method was applied only to 81% of data points. Contrary to cassava, the panel of experts in Zambia was able to identify varieties by names (method E1, Table 12) and type (method E2, Table 13) just by looking at the seed photos.

For beans, experts identified 135 observations of farmer samples by names (method E1) that matched improved varieties, suggesting an adoption rate of 18.3%. The most common name of bean variety reported was Kabulangeti. However, when validating these estimates with DNA results, the adoption rate is much lower at 5.8% because 92 of the 135 observations did not genetically match a released variety, but were mistakenly identified as a released variety by the experts (Table 12).

In contrast, when the panel of experts were asked to identify the farmer seed samples by type (method E2) based on the seed photos, 21% of the data points (142 out of 676) were classified as improved varieties (Table 13).

Table 11. Expert Reported Groupings of Cassava Varieties by Type Based on Panel of Experts' Review of Photos Taken During Visit to Farmers' Field (Method E): Results of the Cassava Varietal Identification Survey, Ghana 2013

Variety type reported by experts (ex post)	Frequency
Improved	142
Local	644
Don't know/missing information	128
Total	914

Table 12. Matching Bean Varietal Names as Identified by Experts Based on Photographs of Seeds Taken by Enumerators during Farmer Surveys (Method E) with Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013

	Variety name as identified by the expert based on seed photos	Variety name in the reference library included for DNA fingerprinting (benchmark)					Sample Does not match any genotype	Total
		Lwan-geni	Lyambai	Lukupa	Kabul-engeti (improved)	Kabul-engeti (local)		
Names of released varieties	Lwangeni	0	0	0	0	1	1	2
	Mbereshi	0	0	0	0	0	0	0
	Kalungu	0	0	0	0	0	0	0
	Lyambai	0	0	0	0	0	4	4
	Kabale	0	0	0	1	0	14	15
	Kapisha	0	0	0	0	0	1	1
	Kabulangeti (improved)	0	0	0	24	8	59	91
	Lukupa	0	0	8	0	0	7	15
	Kalambo	0	0	0	0	0	5	5
	Chambeshi	0	0	1	0	0	1	2
	Kabulengeti (local)	0	0	0	30	7	63	100
	Other Local	2	1	4	48	5	441	501
	Total	2	1	13	103	21	596	736

Table 13. Expert Reported Groupings of Bean Varieties by Type Based on Review of Photos of Seeds (Method E): Results of the Bean Varietal Adoption Survey, Zambia 2013

Variety type reported by experts (ex post)	Frequency
Improved	142
Local	447
Mixture (unclassified)	78
Don't know/missing information	9
Total	676

3.7. Classification of Varieties by Panel of Experts Looking at Samples of Seeds (Beans Only) (Method F)

The panel of experts also looked at seed samples and was asked to provide a consensus variety name (method F1) and type of variety (method F2). Seed samples could not be collected for 109 cases (13%). As can be seen (Table 14), experts identified 123 observations with names of released varieties, suggesting an adoption rate of 14.5%. The most commonly reported name (method F1) was Kabulangeti. However, when validating these results with DNA data, the adoption rate is much lower at 5.2% because 79 of the 123 observations did not genetically match a released variety, suggesting these were mistakenly identified as a released variety (Table 14). When the panel of experts were asked to tell whether the seed sample corresponded to a released variety (method F2), they reported that 122 observations (14.6%) did correspond to this type of variety (Table 15).

Table 14. Matching Bean Varietal Names as Identified by Experts Based on Seed Samples Collected by Enumerators during Farmer Surveys (Method F1) with the Varietal Identification Based on DNA Fingerprinting: Results of the Bean Varietal Adoption Survey, Zambia 2013

	Variety name as identified by the expert based on seed photos	Variety name in the reference library included for DNA fingerprinting (benchmark)					Sample Does not match any genotype	Total
		Lwangeni	Lyambai	Lukupa	Kabul-engeti (improved)	Kabulengeti (local)		
Names of released varieties	Lwangeni	0	0	1	0	0	7	8
	Mbereshi	0	0	0	0	0	0	0
	Kalungu	0	0	0	0	0	0	0
	Lyambai	0	0	0	0	0	3	3
	Kabale	0	0	0	2	0	15	17
	Kapisha	0	0	0	0	0	0	0
	Kabulangeti (improved)	0	0	0	33	3	38	74
	Lukupa	0	0	8	0	0	5	13
	Kalambo	0	0	0	0	0	6	6
	Chambeshi	0	0	0	0	0	2	2
	Kabulengeti (local)	0	0	0	28	11	93	132
	Other Local	3	1	4	58	9	517	592
	Total	3	1	13	121	23	686	847

Table 15. Expert Reported Groupings of Bean Varieties by Type Based on Panel of Experts' Review of Seed Samples (Method F2): Results of the Bean Varietal Adoption Survey, Zambia 2013

Variety type reported by experts (ex post)	Frequency
Improved	122
Local	573
Mixture (unclassified)	103
Don't know/missing information	37
Total	835

3.8. Validating Results of Individual Methods against DNA Fingerprinting

As the results suggest, each method evaluated provides different estimates of adoption rates of improved varieties. This raises serious questions about which method is the most accurate or closer to the truth in estimating varietal adoption in farmers' fields. We address this question by validating the results of varietal identification based on all these methods against the benchmark of DNA fingerprinting. This comparison of results of varietal identification based on these different methods against the DNA fingerprinting results are presented in Tables 16-18.

Table 16. Measures of Effectiveness of Different Methods of Varietal Identification for Cassava in Ghana: Comparison of Outcomes against the Benchmark of DNA Fingerprinting Under the *Liberal* Assumption

Measures of effectiveness		Method A1 *Farmer reported name*	Method A2 *Farmer reported type*	Method D2a *Expert reported name based on field observation*	Method D2b *Expert reported type based on field observation*	Method E2 *Expert group opinion on type based on photos*	Benchmark (DNA)
Varietal data points compared	N	914	914	914	914	914	914
Outcome: Number of data points classified as 'released' (or improved) varieties	N	20	51	33	47	142	263
	%	2%	6%	4%	5%	16%	29%
Type I Error A local variety incorrectly identified as released variety	N	0	14	0	17	104	
	%		27%		36%	74%	
Type II Error A released variety incorrectly identified as a local variety or by an incorrect RV name	N	254	226	263	233	193	
	%	97%	86%	92%	89%	73%	
Accuracy of name: Data points correctly identified as RV by name	N	9	--	20	--	--	
	%	3%		8%			
Accuracy of category: Data points correctly identified as RV by Type	N	--	37	--	30	70	
	%		14%		11%	27%	

Under the liberal scenario for cassava, the DNA results suggest that 29% of the farmer samples corresponded to improved varieties (Table 16). From the five methods evaluated against this benchmark, none provides an estimate close to this value. The estimates using farmer based elicitation approaches or expert based field observations or photo based identification ranged from 2% (method A1) to 16% (method E), which suggests that any estimates of adoption of improved varieties would be underestimated using these methods. However, in all methods, either local varieties were incorrectly identified as IVs (Type I error) or improved varieties were incorrectly identified as local variety or by an incorrect IV name (Type II error). The most important results are provided in the last two rows of the table, as these indicate how many data points or observations were correctly identified as IVs either by name or type, when validated against the DNA results.

As one might expect, the results are very different under the conservative scenario for cassava, when cluster groups 2 and 3 based on the DNA results are classified as local varieties (see Table 3). In this scenario, only 4% of farmer samples are considered same as released varieties (Table 17). Against this benchmark and definition of released versus local varieties, two methods provide estimates close to this value: method A2 (6%) and method

D2b (5%). On the other hand, method E greatly overestimates adoption of IVs when compared with this conservative way of interpreting the DNA results. Similar to the liberal scenario, Type I and Type II errors were observed in all methods. In this conservative scenario, the method that proved most accurate to identify IVs was method A2 (i.e., using the type of variety reported by farmers), as 55% of the variety type reported were correctly identified as IV (i.e., matched DNA results). Similar to the liberal scenario, the results suggest that varietal identification by name using these different farmer and expert based methods is more difficult than identifying variety as improved or local.

For beans, the DNA analysis included extra samples because for those varieties reported as mixtures by the farmers, the seeds were separated by their color (or market class) for DNA analysis. This generated additional data points than the number of observations based on household surveys. This also explains the difference in the number of data points between Table 13 and Table 18 for method E.

Table 17. Measures of Effectiveness of Different Methods of Varietal Identification for Cassava in Ghana: Comparison of Outcomes against the Benchmark of DNA Fingerprinting under the *Conservative* Assumption

Measures of Effectiveness		Method A1 *Farmer reported name*	Method A2 *Farmer reported type*	Method D2a *Expert reported name based on field observation*	Method D2b *Expert reported type based on field observation*	Method E2 *Expert group opinion on type based on photos*	Benchmark (DNA)
Varietal data points compared	N	914	914	914	914	914	914
Outcome: Number of data points classified as 'released' (or improved) varieties	N	20	51	33	47	142	40
	%	2%	6%	4%	5%	16%	4%
Type I Error A local variety incorrectly identified as IV	N	4	29	4	28	121	
	%	44%	57%	20%	60%	85%	
Type II Error An released variety incorrectly identified as a local variety or by an incorrect IV name	N	35	18	27	21	21	
	%	88%	45%	68%	53%	53%	
Accuracy of name: Data points correctly identified as IV by name	N	5	--	13	--	--	
	%	13%		33%			
Accuracy of category: Data points correctly identified as IV by Type	N	--	22	--	19	19	
	%		55%		48%	48%	

The results suggest that based on DNA fingerprinting of farmer samples and comparing those fingerprints against the fingerprints of released varieties in Zambia, 16% of the farmer samples correspond to released varieties (Table 18). From the five methods evaluated against this benchmark, three methods provided an estimate close to this value: method A2 (13%), method E (18%) and method F (15%). Method C gave a gross overestimation of adoption of IVs. Similar to cassava, Type I and Type II errors were present in all the methods. The results also suggest that the method traditionally used to estimate adoption of improved varieties (i.e., farmer survey based methods A1 and A2) could potentially provide misleading results, since only 9% of the names reported by farmers accurately matched the variety as identified by DNA results.

Table 18. Measures of Effectiveness of Different Methods of Varietal Identification for Beans in Zambia: Comparison of Outcomes against the Benchmark of DNA Fingerprinting

Measures of Effectiveness		Method A1 *Farmer reported name*	Method A2 *Farmer reported type*	Method C *Farmer identification based on matching seed samples*	Method E *Expert group opinion based on photos*	Method F *Expert group opinion based on seed sample*	Benchmark (DNA)
Varietal data points compared	N	855	802	783	736	847	855
Outcome: Number of data points classified as 'released' (or improved) varieties per the given Method	N	34	106	553	135	123	141
	%	4%	13%	71%	18%	15%	16%
Type I Error A local variety incorrectly identified as IV	N	21	85	441	101	84	
	%	62%	80%	80%	75%	68%	
Type II Error An released variety incorrectly identified as a local variety or by an incorrect IV name	N	129	108	59	87	97	
	%	91%	84%	48%	73%	70%	
Accuracy of name: Data points correctly identified as IV by name	N	12	--	63	32	41	
	%	9%		52%	27%	29%	
Accuracy of category: Data points correctly identified as IV by Type	N	--	21	--	*34*	*44*	
	%		16%		29%	32%	

4. CONCLUSIONS AND IMPLICATIONS

4.1. Key Results and Lessons Learned

The two studies described in this paper represent one of the first pilots to test different methods of varietal identification and validation against DNA fingerprinting. The main purpose of these pilot studies was to find out which method is able to accurately identify released varieties. The rationale behind this pilot was to come up with lessons learned and recommendations on methods/approaches that can be used in scaling up the collection and assembly of diffusion data on improved varieties, either as part of routinely conducted nationally representative household level surveys or by seeking opportunities to institutionalize a simple and cost-effective approach that can be implemented by the research program to track varietal adoption. We highlight key results and identify several lessons learned from these pilot studies that can help guide future work in this area.

- One of the key results of this study is that for both cassava and beans, the individual methods studied provided different estimates of adoption of improved varieties, ranging from 2% to 16% for cassava and 4% to 71% for beans. Moreover, no one method stood out to be most effective in identifying varieties accurately on all measures of effectiveness. Thus, we are not able to recommend any method tested in these two pilots as an identification method for tracking varietal adoption.

- Secondly, some of the methods proposed for cassava that required recording morphological traits of plants at different stages of plant growth proved difficult to implement because of their operational complexity or problems in soliciting responses from the farmers or experts. The scalability of such methods at a nationally representative level is thus questionable.[3]

- Third, at an aggregate level, some methods provided estimates of adoption closer to the truth (as determined by the DNA results) than others. However, at an individual observation level, all methods suffered from both type I (i.e., local varieties were incorrectly identified as improved varieties) and type II errors (i.e., improved varieties incorrectly identified as improved varieties). These results, thus, raise questions on the credibility of estimating varietal adoption based on the so-called gold standard of eliciting this information from the farmers as part of the household surveys. It also raises questions on the accuracy of any analysis that uses adoption estimates derived from such methods as an outcome variable (as is done in the determinant of adoption studies) or as one of the independent variables (as done in impact studies that try to estimate the causal link between varietal adoption and welfare outcomes). How significant are these potential errors in such analyses needs to be investigated and a topic of future research.

- Fourth, identifying farmer grown improved varieties accurately by name in a setting where there is a diversity of names by which farmers call their varieties and the seed system is predominantly informal, is a challenge across all the methods tested. Results for both beans and cassava indicate that in such a setting the traditional method of farmer elicitation will give an underestimate of adoption of improved varieties.

[3] It is worth noting that there are efforts underway to improve phenotyping methods using imagery, remote sensing, and various forms of spectroscopy that may provide alternative methods of variety identification in the near future.

- Fifth, farmers and experts were generally better able to give an aggregate assessment of the adoption of improved varieties as a category than variety specific adoption. But methods that can only identify varieties by the category of improved versus traditional variety are not useful in assessments of varietal turnover, and in estimating type II benefits of plant breeding research that are generated when farmers replace older varieties with new generation of improved varieties (Morris and Heisey 2003). Therefore, the potential of these methods that can only provide aggregate level adoption estimates of varieties by category or type may be limited in scope.

4.2. DNA Fingerprinting as a Method of Varietal Identification: Potential and Limitations

One of the main results and conclusions derived from these two pilots is that for the purposes of tracking variety specific adoption and varietal turnover in farmers' fields, the only credible method is DNA fingerprinting. This method provides a true picture of what is in farmers' fields as compared to a reference library of released varieties and other landraces. It also provides a true picture of the genetic identity of released varieties included in the reference library, which sometimes may not conform to expectations and can pose a challenge in classifying varieties grown by farmers. For example, as shown by the case of cassava, DNA analysis revealed that multiple released varieties were essentially identical; and some accessions considered to be released varieties shared the same DNA fingerprints as accessions considered to be landraces. Because of this unexpected finding in the case of cassava, identification of farmer adopted varieties by name or by type had to be assessed under two scenarios—liberal and conservative. This undermined the purpose of this study, which was to identify farmer samples with specific and unique released varieties.

An important lesson learned from the cassava pilot study is that DNA-based variety identification is only as good as the quality of the library. If the library is inadequate, not curated, or contains misidentified and/or redundant accessions which bear different names, then the outcome of the work will not be reliable. The results of this study, thus, point to the importance of first making sure the released varieties included in the reference library have unique genetic identities. If they are found not to be genetically unique, this information can help guide the methodology of the adoption study and help make the decision on the utility and value of doing DNA fingerprinting at the farm level. Such information can also help further investigate the functioning of the varietal development, release and dissemination system for that crop and country and help guide the policies governing the varietal release and seed system.

The two pilot studies reported in this paper also point to several challenges and lessons learned that need to be considered in scaling up DNA fingerprinting-based method for varietal identification as part of household surveys. First, there are logistical challenges of collecting, tracking, storing, and transporting the samples from farmers' fields to a lab facility to get high quality DNA. The experience of these two pilots suggests that proper training of the field staff and access to resources to make sure the logistics are in place to maintain the sample identity and quality from the farm to the lab is very critical. Second, the logistical challenges of sample collection associated with a vegetative crop such as cassava were different from a grain crop like beans. To get high quality DNA for some crops may require visiting farmers or their fields at certain time of the growing season, which may not correspond with the planned household surveys. This can potentially increase the field costs of the overall study, if the collection of samples for DNA analysis will require additional visits to farm households or their fields as was the case with cassava. If the plan is to collect

seed samples, this can be done as part of the household survey and, thus, lower the overall cost. But the window of opportunity to do the survey may be quite narrow to ensure farmers will have some seeds to share at the time of the survey. As we experienced in Zambia, many farmers had no beans left to share with the team even one month after the harvest when the survey was done.

Third, tracking varietal adoption based on integrating DNA fingerprinting as part of household level surveys poses the additional challenge of sampling to make sure the samples subjected to DNA analysis are representative of the varieties grown by the household that is selected for the representative survey. Issues such as number of samples to be collected per household, method of sample selection that is representative of all farmer grown varieties, and number of plots and plants from which samples should be collected, type and timing of sample collection needs careful consideration. The field staff needs to be well trained in operationalizing that sampling plan to make sure the DNA fingerprinting results are representative of the varieties grown by the selected households.

Lastly, the scaling up of DNA fingerprinting as part of representative household surveys will depend on the cost, which includes not only the marginal cost of field sample collection, but also the cost of establishing the reference library, DNA extraction, genotyping, and data analytics. In this study, the total estimated cost per data point for these four cost categories (i.e., reference library, DNA extraction, genotyping, and data analytics) was estimated to be $42 for cassava and $35 for beans. The cost of genotyping as a percentage of total cost was about 75% in the case of cassava and 66% in the case of beans. These differences in the cost of genotyping between the two crops are mainly due to different techniques used for DNA fingerprinting. The GBS method used for cassava is relatively more expensive and generates large amounts of data (all of which may not be needed for varietal differentiation) than the two-step SNP marker technique used for beans. There are also other types of DNA fingerprinting techniques (e.g., Diversity Arrays Technology and SSR) that can potentially be used for varietal identification, each with different cost implications, type of sample requirement, and types of information that can be generated.

It is important to note that this study demonstrated the value of using DNA fingerprinting with the current available methods. Overall, the methods are improving rapidly and the cost of genotyping across these different methods is rapidly declining.[4] Thus, the total cost of high volume DNA fingerprinting per sample are projected to come down. However, unless the cost per data point is reduced substantially from the cost experienced in these two pilots, the use of DNA fingerprinting as part of large-scale representative surveys may be long way from becoming routine. One suggestion to consider for reducing the cost and to make the logistics more manageable would be to use DNA fingerprinting as a method of validation on a random sub-sample of households rather than all the households in a large-scale representative surveys. Expanding the capacity to do high volume DNA fingerprinting within the country or easy access to such capacity regionally/internationally, with no government restrictions on the shipment of plant tissues or DNA samples to other countries, are other potential ways to reduce the cost of DNA fingerprinting and should be considered for future efforts.

Finally, we would like to add this caveat that the lessons highlighted and recommendations made on the potential use of DNA fingerprinting for varietal identification are derived from

[4] For example, the GBS method used for cassava gave data on 56,000 SNPs at the cost of about $28/sample. For the purpose of varietal identification, one does not need large amounts of data on thousands of SNP markers. Scientists are now working on a technology that will give data for up to 5000 SNPs at < $10/sample, which should be quite adequate to distinguish varieties with reasonable level of accuracy.

the experience base of only two case studies described in this Report. More studies on different crops and country settings are needed to generate an experience base and derive generalizable conclusions.

ANNEX

Annex 1. List of Cassava Released Varieties Included in the Reference Library for Ghana

s/n	Variety Name \a	Year released	Institutional Origin/Final Cross/Country	Variety Attributes/Pedigree
1	Abasafitaa	1991	IITA/CRI	Yield, pest, and disease resistant
2	Afisiafi	1993	IITA/CRI	Yield, pest, and disease resistant
3	Tek Bankye	1997	IITA/KNUST	High yield, good cooking quality
4	Nyerikobga	2002	SARI	Lower gari swelling power
5	Filindiakong	2002	IITA/CRI	Considered early, high DM
6	IFAD	2004	KNUST	High yield, good cooking quality
7	Nkabom	2004	KNUST	Poundability, high yield
8	UCC (Capevars bankye)	2005	UNIV OF CAPE COAST	Poundability, high yield
9	Bankye botan	2005	UNIV OF CAPE COAST	Poundability, high yield
10	Bankyehemaa	2005	IITA/CRI	Poundability, high yield
11	Esam bankye	2005	IITA/CRI	High yield/dry matter, flour
12	Dokuduade	2005	IITA/CRI	High yield, starch content
13	Agbelifia	2005	IITA/CRI	High yield/dry matter, flour
14	Otuhia	2009	IITA/CRI	High yield disease/pest tolerance
15	Sikabankye (Sika)	2009	IITA/CRI	High yield, starch content
16	Bankye broni \b	2009	IITA/CRI	High yield disease/pest tolerance
17	Ampong	2009	IITA/CRI	High yield disease/pest tolerance

\a two released varieties – Gblemoduade, released in 1993 and Eskamaye, released in 2002 were not included in the reference library as they were not physically available at the CRI research station for sample collection.
\b Two copies of Bankye _broni were included in the reference library, which makes the total number of released variety samples = 18

Annex 2. List of Cassava Landraces Included in the Reference Library for Ghana

N	Variety Name	N	Variety Name	N	Variety Name	N	Variety Name
1	12_0197	13	AW3_10_011	25	DMA2000_66	37	OFF_2000_023
2	12_0236	14	BNSIA_TUMTUM	26	DWA2000_069	38	OFF_2000_037
3	12_02Y5	15	BOSOMENSIA_1	27	ELISHA	39	OFF_2000_133
4	ABUSUA	16	BOSOMENSIA_2	28	ESIABAYAA	40	UCC2000_110
5	ADE2000_182	17	BRONI	29	KSI2000_125	41	UCC20001_464
6	ADW2000_002	18	CONGO_BATIALION	30	KW_2000_009	42	UCC2001_249
7	ADW2000_003	19	DEBOR_1	31	KW_2000_030	43	UCC2001_399
8	ADW2001_051	20	DEBOR_3	32	KW_2000_148	44	UCC2001_448
9	AFS_2000_050	21	DEBOR_BEPOSO	33	KW2000_181	45	WCH2000_010
10	ANKRA	22	DEBOR_KAAN	34	KWANWOMA	46	WCH2000_019
11	ANKRA_10_003	23	DMA2000_002	35	MONICA		
12	AW3_10_008	24	DMA2000_031	36	OFF_2000_018		

Annex 3. List of Bean Released Varieties Included in the Reference Library for Zambia

Variety name	Year of release	Seed color	Seed size	Growing Habit	Color of flowers	Color of immature pods	Color of mature pods (before drying)
Chambeshi	1998	Khaki/ cream	Large	Bush	White	Green	Cream
Lukupa	1999	Cream mottled	Medium	Indeterminate bush	White	Green	Cream with stripes
Lyambai	1999	Red mottled	Medium	Bush	Pinkish	Green	Cream
Kalungu	2004	White	Medium	Indeterminate bush	White	Green	Cream
Kabulangeti	2007	Purple	Medium	Indeterminate (semi-climber)	White	Green	Cream with purplish stripes
Kapisha	2007	Cream	Medium	Indeterminate (semi-climber)	White	Green	Cream
Kabale	2007	Pinkish	Medium	Bush	Pinkish	Green	Cream
Lwangeni	2009	White	Small	Indeterminate bush	White	Green	Cream
Kalambo	2011	Cream mottled	Large	Indeterminate Dwarf	White	Green with speckles	Cream with speckles
Sadzu	2011	Red mottled	Large	Climber	White	Green	Cream
Mbereshi	2012	Red mottled	Large	Bush	White	Green	Cream

Annex 4. List of Farmer Reported Cassava Varieties Identified by them as Improved Varieties

Variety name\a	Frequency
AGRIC	18
BANKYE_HEMAA	4
NKABOM	4
KUFFOUR	3
DEBOR	2
ESI	2
MEDUMEKU	2
MPOHOR	2
ABASA_FITAA	1
ABOSOMNSIA	1
AFISIAFI	1
AKOROFUOMPE	1
AMENFI	1
AMPONG	1
ANKRA	1
BANKYE_FITAA	1
BOSOMENSIA	1
ESAM_BANKYE	1
ESIABAYAA	1
GBEZEH	1
SANTOM_BANKYE	1
SIKA_BANKYE	1
Total	51

\a Variety names that match the released varieties in Ghana are bolded.

Annex 5. List of Farmer Reported Bean Varieties Identified by them as Improved Varieties

Variety name\a	Frequency
Kabulangeti	37
Lusaka	18
White	15
Pomba	9
Solwezi	9
Red Beans	4
Buteko	3
Kapwepwe	2
Kontola	2
Sweet Beans	2
White/Yellow	2
Yellow	2
Fungula	1
Imiti Ikula	1
Kamuti	1
Kapisha	1
Lukupa	1
Lwangeni	1
Miti	1
Mixed	1
Miyombe	1
Nakambalala	1
Nambalala	1
Nyombe	1
Pamba	1
Simpilila	1
Special	1
Tamba	1
Don't know	1
Total	122

\a Variety names that match the released varieties in Zambia are bolded.

Annex 6. List of Expert Reported Variety Names Identified by them as Improved Varieties

Variety name \a	Frequency
Bankye Hemaa	11
Afisiafi	9
Nkabom Bankye	9
No name	8
Abasafitaa	2
Esam Bankye	2
Agric	2
Bosomensia	1
Lagos	1
Antifo Bankye	1
Sika	1
Total	47

\a Variety names that match the released varieties in Ghana are bolded.

REFERENCES

Alene, A.D., A. Menkir, S.O. Ajala, A.S. Badu-Apraku, V. Olanrewaju, M. Manyong, and A. Ndiayel. 2009. The Economic and Poverty Impacts of Maize Research in West and Central Africa. *Agricultural Economics* 40.5: 535–550.

Angelucci, F. 2013. Analysis of Incentives and Disincentives for Cassava in Ghana. FAO Technical Notes Series. Rome: MAFAP, FAO.

Evenson, R.E. and D. Gollin (Eds). 2003. Crop Variety Improvement and its Effect on Productivity: The Impact of International Agricultural Research. Oxon, UK: CABI.

FAOSTAT. 2014. Statistical Database. Rome, Italy: FAO.

Floro, IV, O. Victorino, Ricardo A. Labarta, Luis A. Becerra López-Lavalle, Jose M. Martinez, and Tatiana M. Ovalle. 2016. Household Determinants of the Adoption of Improved Cassava Varieties Using DNA Fingerprinting to Identify Varieties in Farmer Fields: A Case Study in Colombia. Draft paper.

Griliches, Z. 1958. Research Costs and Social Returns: Hybrid Corn and Related Innovations. *Journal of Political Economy* 66.5: 419–431.

Kassie, M., B. Shiferaw, and G. Muricho. 2011. Agricultural Technology, Crop Income, and Poverty Alleviation in Uganda. *World Development* 39.10: 1784–1795.

Kosmowski, Frederic, Abiyo Aragaw, Andrzej Kilian, Alemayehu A. Ambel, John Ilukor, Biratu Yigezu, and James Stevenson.2016. Varietal Identification in Household Surveys: Results from an Experiment Using DNA Fingerprinting of Sweet Potato Leaves in Southern Ethiopia. Policy Research Working Paper No. 7812. Washington, DC: The World Bank.

Mathenge, Mary K., Melinda Smale, and John Olwande. 2014. The Impacts of Hybrid Maize Seed on the Welfare of Farming Households in Kenya. *Food Policy* 44: 262–271.

Morris, M.L. and P.W. Heisey. 2003. Estimating the Benefits of Plant Breeding Research: Methodological Issues and Practical Challenges. *Agricultural Economics* 29: 241–252.

Raatz, Bodo. 2015. Analysis of SNP Data of Samples Collected in Zambia to Quantify Impact of Variety Releases. Mimeo.

Rabbi, Ismail, Peter Kulakow, Joseph Manu-Aduening, Ansong Dankyi, James Asibuo, Elizabeth Parkes, Tahirou Abdoulaye, Gezahegn Girma, Melaku Gedil, Punna Ramu, Byron Reyes, and Mywish Maredia. 2015. Tracking Crop Varieties Using Genotyping-By-Sequencing Markers: A Case Study Using Cassava (*Manihot esculenta* Crantz). *BMC Genetics* 16: 115. DOI 10.1186/s12863-015-0273-1.

Shiferaw, Bekele, Menale Kassie, Moti Jaleta, and Chilot Yirga. 2014. Adoption of Improved Wheat Varieties and Impacts on Household Food Security in Ethiopia. *Food Policy* 44: 272–284.

Walker, T.S. and J. Alwang. (Eds). 2015. I*mproved Varieties of Crops in Sub-Saharan Africa: Assessing Progress in the Generation, Adoption and Impact of New Technologies.* Montpellier, France and Wallingford, U.K.: *CGIAR and* CAB International.

Yirga, C., G. Traxler, M. Kim, and D. Alemu. 2015. Using DNA Fingerprinting to Estimate the Bias of Farm Survey Identification of the Diffusion of Improved Crop Varieties in Ethiopia. Paper presented at the International Conference of Agricultural Economists 9-14 August. Milan.

Zeng, Di, Jeffrey Alwang, George W. Norton, Bekele Shiferaw, Moti Jaleta, and Chilot Yirga. 2015. Ex Post Impacts of Improved Maize Varieties on Poverty in Rural Ethiopia. *Agricultural Economics* 46.4: 55–526. DOI: 10.1111/agec.12178.