Nigeria Country Plan Baseline and Varietal Monitoring Survey

DNA Based Varietal Identification of Cowpea Varieties in Nigeria

By

Tesfamicheal Wossen, Ismail Rabbi, Bamire, S. Adebayo, Unachukwu Nnanna, Boukar Ousmane, Melaku Gedi, Julius Manda, Chiristian Fatokun, Arega Alene, Tahirou Abdoulaye and Victor Manyong.

Prepared for the Bill and Melinda Gates Foundation

Published by
Obafemi Awolowo University Press
Ile-Ife, Nigeria

ISBN: 978-136-323-1

Printed in Nigeria by:
Obafemi Awolowo University Press
Ile-Ife, Nigeria.

Recommended Citation:
Tesfamicheal Wossen, Ismail Rabbi, Bamire, S. Adebayo, Unachukwu Nnanna, Boukar Ousmane, Melaku Gedi, Julius Manda, Chiristian Fatokun, Arega Alene, Tahirou Abdoulaye and Victor Manyong. 2019. Nigeria Country Plan Baseline and Varietal Monitoring Survey DNA Based Varietal Identification of Cowpea Varieties in Nigeria, Obafemi Awolowo University Press, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria.

.

Contact:
Professor Adebayo Simeon Bamire
Department of Agricultural Economics, Faculty of Agriculture
Obafemi Awolowo University
Ile-Ife, Osun State, Nigeria
Mobile: +234 8033898900
Email: abamire@oauife.edu.ng

And

Dr. Tahirou Abdoulaye
International Institute of Tropical Agriculture (IITA),
PMB 5320, Oyo Road, Ibadan 200001,
Oyo State, Nigeria
Mobile: +234 8039784465
Email: T.Abdoulaye@cgiar.org

This report was prepared by:
International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria

Table of Contents

List of Tables

List of Figures

Acknowledgements

The authors would like to gratefully acknowledge the support and cooperation provided by institutions, individuals, and other stakeholders. The authors are especially grateful to:

- The Bill and Melinda Gates Foundation (BMGF), who provided funding for this study
- Dr. Hannah Reed of the BMGF for her immense hands-on support throughout the project. The initial Program officer for this project was Elvis Fraser and Dr. Hannah Reed took over the oversight after his departure from the foundation. We are immensely grateful to both of them (and the entire BMGF team) for the coaching and support that led to this output.
- The Nigeria National Bureau of Statistics (NBS), who provided technical assistance towards data collection
- Mrs. Olabisi Olajire and the entire staff of the National Centre for Genetic Resources and Biotechnology (NACGRAB)
- The Federal Ministry of Agriculture and Rural Development (FMARD), who provided technical support towards data collection
- Dr. Melaku Gedil and other IITA scientists for their insights about maize.
- All supervisors, enumerators and drivers who participated in the field activities. Their dedication, meticulous attention to detail, and flexibility throughout the training and fieldwork are unparalleled. We would also like to thank various state and local governments' Agricultural Development programs, along with the National Bureau of Statistics officers who assisted us in reaching a random sample of communities and rural farming households. Their commitment towards collecting reliable data is highly appreciated.

Finally, special thanks go to the rural farming households, extension agents and local leaders across the project areas for graciously assisting the scientists and researchers from OAU, IITA and NISER. We enjoyed learning and working together to transform Nigeria's agriculture.

List of Acronyms

ADP	Agricultural Development Program
AWA	Area Weighted Average Age
A-WEAI	Abbreviated Women's Empowerment in Agriculture Index
BMGF	Bill and Melinda Gates Foundation
DNA	Deoxyribonucleic Acid
FMARD	Federal Ministry of Agriculture and Rural Development
GPS	Global Positioning System
IAR	Institute of Agricultural Research
IAR&T	Institute of Agricultural Research & Training
ID	Identification
IITA	International Institute of Tropical Agriculture
NACGRAB	National Centre for Genetic Resources and Biotechnology
NBS	National Bureau of Statistics
NIBAS	Nigeria Baseline Study
NISER	Nigerian Institute of Social and Economic Research
OAU	Obafemi Awolowo University
OLS	Ordinary Least Square
SNP	Single Nucleotide Polymorphism
WTA	Willingness To Accept

Executive Summary

The development and dissemination of improved crop varieties is the primary pathway through which technological change in the agricultural sector can bring about productivity gains. Despite the substantial investment in breeding that led to the development and dissemination of more than 30 improved cowpea varieties in Nigeria, adoption rates of improved varieties have not been well documented. The Nigeria Baseline Study (NIBAS) project was designed to fill this gap by generating data to measure adoption rates across six states (Benue, Kaduna, Kano, Katsina, Nasarawa, and Niger) in Nigeria. This report therefore documents key evidence on the adoption of improved cowpea varieties across the six states of Nigeria using two approaches: (i) a traditional survey approach in which information about improved varieties was directly elicited from farmers, and (ii) a DNA fingerprinting approach which helps to accurately identify the cowpea varieties grown by farmers. The report will focus on documenting the process of DNA-fingerprinting, including genotype collection at farmers field, rates and intensity of adoption, as well as the characterization of the most widely adopted improved varieties, in terms of area coverage, geographical distribution and age. The study sampled households from the six NIBAS states representing the largest cowpea production areas in Nigeria. The analysis was based on 1497 different cowpea leaf samples collected from 1243 different farmers. The results obtained are highlighted as follows:

Varietal Identification

- The traditional variety identification methods for tracking adoption of improved varieties include: secondary sources (e.g., published government reports, available data sets); seed sales and seed multiplication/distribution data; expert opinion/key informant interviews; Community level surveys including focus group discussion; farmer elicitation as part of a household survey; and use of morphological descriptors. However, DNA fingerprinting offers an alternative approach in which samples from farmers' fields can be matched to known improved varieties and landraces.
- A well curated comprehensive reference library was developed in collaboration with breeding programs for tracking genotypes of interest. Collections representing known farmer-preferred landraces were included in the reference library to identify common or popular unimproved varieties. The seeds for the reference library were received from breeding organisations such as the National Centre for Genetic Resources and Biotechnology (NACGRAB), the Nigeria National Gene Bank, and Seed Companies. In the absence of a useful library, initial DNA fingerprinting served to establish a baseline data.
- The most common improved variety uncovered in the study was IT89KD-374. The second most common improved variety was IT99K-216-24-2. Other identified improved

varieties were IAR353, SAMPEA_2, IT98K-131-2, IT89KD-288 and IT81D-985. The remaining 12 varieties occurred with less frequency.

Rate and Intensity of Adoption

- Adoption rate as measured by DNA-based varietal identification was 24.4% based on the assumption that all the unidentified/unknown varieties are unimproved.
- Adoption rate was considerably low with variations across all states. Survey data showed the adoption rate ranging from 11.6% in Katsina to 42.2% in Kano while DNA fingerprinting showed a range from 10.8% in Kaduna to 36.3% in Kano.
- The intensity of adoption, measured by the total cowpea area under improved cowpea varieties was 16.3% based on survey data, and 17% based on DNA-fingerprinting, implying that at least 17% of the cowpea area in the NIBAS sample is under improved varieties.

Geographical Spread of Adopted Improved Varieties

- Majority of the common varieties were found across a large part of Northern Nigeria, from Kebbi in the East to Gombe and Borno in the East, suggesting that these varieties are adapted to the regions that share common agroecological characteristics. Some varieties, particularly, the less common genotypes, were found in smaller geographical areas.
- The area weighted average age of improved varieties in the NIBAS sample was 20.2 years, implying that, on average, it takes about 20 years for a given variety to be replaced by another variety.

Classification of Seed Types and Sources

- Only 6.3% of the farmers correctly identified the improvement status of the cowpea varieties they planted; 15.8% misidentified improved varieties as local, believing that they were growing local varieties but actually growing improved varieties; and 2.6% of the farmers misidentified local varieties as improved, believing that they were growing improved varieties but actually growing local varieties.
- There was no correlation between seed sources and variety identification while a positive correlation was established between adoption of improved cowpea varieties and productivity. Thus, agricultural development programs and policies should address both the quantity of adoption (low adoption rates) and the quality of adoption. This requires efforts beyond mere dissemination of improved varieties to achieve productivity growth and calls for quality control and dissemination of highly quality seeds as an integral part of the dissemination process. There is also a need for further analysis to understand the nature of the "unknown/unidentified" varieties so as to disseminate newer and better improved varieties in the areas where these varieties are popular.

1. Background

It is widely recognized that the development and dissemination of improved crop varieties is the primary pathway through which technological change in the agricultural sector can bring about productivity gains. In this regard, crop improvement programs in Nigeria, the largest producer and consumer of cowpea in the world, have focused on development of improved cowpea varieties that are not only high yielding under optimal conditions, but which can also overcome stresses from *Striga* infestation and drought. This investment has led to the development and release of more than 30 improved cowpea varieties since 1980's (Figure 1).

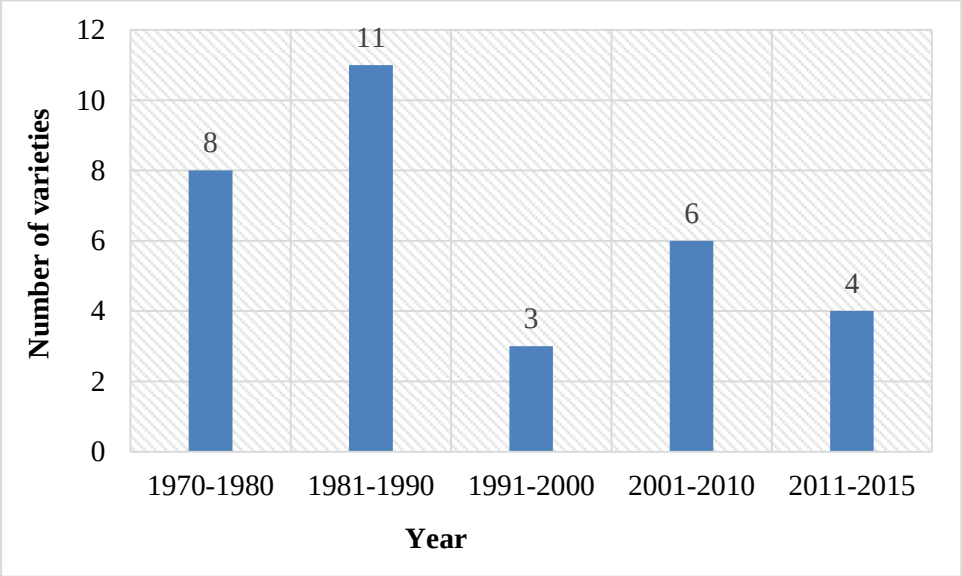

Figure 1: Summary of released cowpea varieties in Nigeria

Despite the substantial investment in breeding that led to the development and dissemination of more than 30 improved cowpea varieties in Nigeria, adoption rates of improved varieties have not been well documented. The Nigeria Baseline Study (NIBAS) project was designed to fill this gap by generating data to measure adoption rates across six states (Benue, Kaduna, Kano, Katsina, Nasarawa, and Niger) of Nigeria as shown in Figure 2.

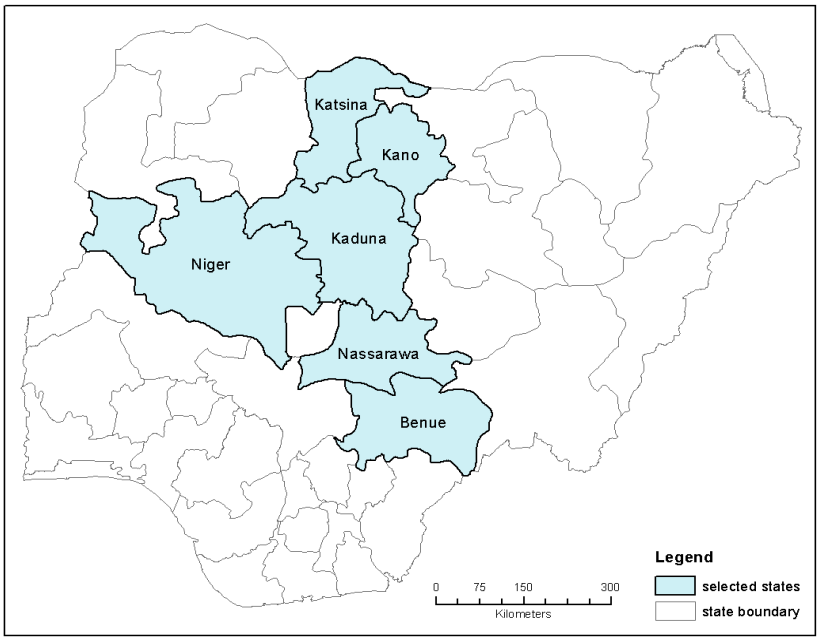

Figure 2: Map of Nigeria showing the six states included in NIBAS
Source: IITA GIS unit, 2019

The NIBAS project collected comprehensive information on the livelihoods and agricultural behaviors and outcomes in the above six states using (i) large-scale household survey, (ii) DNA-based identification of improved varieties, and (iii) gender empowerment indicators using Abbreviated Women's Empowerment in Agriculture Index (A-WEAI).

In this report, we present key evidence on the adoption of improved cowpea varieties. Since accurate measurement and verification of the authenticity of improved varieties is critical, this project employed a DNA-fingerprinting approach to accurately identify the improvement status of the cowpea varieties grown by farmers. DNA fingerprinting offers a reliable method to accurately identify varieties grown by farmers and serves as a benchmark against which to compare the effectiveness of other potential methods for scaling up. As such, we report the adoption rate of improved varieties across the six states of Nigeria using two approaches. The first approach is based on the traditional survey approach in which information about improved varieties was directly elicited from farmers. The second approach is through DNA fingerprinting, which helps to accurately identify the cowpea varieties grown by farmers.

The report will focus on documenting the process of DNA-fingerprinting, including the genotype collection at farmers field, documenting adoption rates, characterization of the most widely

adopted improved varieties, in terms of area coverage, geographical distribution and age. The analysis is based on 1497 different cowpea leaf samples collected from 1243 different farmers. For the purpose of genotyping, we collected cowpea leaf samples from the main cowpea plots of a specific cowpea grower. Field sampling strategy for DNA fingerprinting depends on reproductive biology of a species which can vary from fully outcrossing to self-pollinating as well as propagation of planting materials (seeds or clones). Cowpea is primarily self pollinating with relatively low levels of within individual heterozygosity and within variety genetic diversity. Thus, a single plant/seed is sufficient to represent genotype profile of a variety if it is true-to-type. However, to capture potential within-field variation due to rare outcrossing events and/or seed mixtures, we decided to sample 12 plants per variety per field for bulked DNA extraction and sequencing. In total, leaf samples in more than 1800 plots were collected from farmers field. However, some samples were lost in the process of data collection and genotyping. Hence, our final genotyped sample size comprises of different leaf samples from 1497 plots. Table 1 presents the geographical distribution of the collected samples across the six states. The number of samples from each state ranged from 117 in Nasarawa to 408 in Katsina, with the exception of Benue where only 41 samples were collected for DNA fingerprinting. Number of varieties cultivated by the different households in the NIBAS survey ranged from 1 to 5 with the majority growing single or two cultivars.

Table 1. Total number of samples collected for DNA-fingerprinting

	Samples collected (#)	Cowpea farmers samples were collected (#)	Total cowpea growing farmers in the sample (#)	Share (%)
Full sample	1497	1243	1738	71.5
Nasarawa	117	106	119	89
Benue	41	28	164	17
Kaduna	360	319	348	91.6
Niger	217	176	248	70.9
Kano	354	276	365	75.6
Katsina	408	338	494	68.4

2. DNA-based varietal identification using Single Nucleotide Polymorphisms (SNP)

2.1 Introduction

Impact assessment of crop improvement programs requires accurate measurement of adoption rates of improved varieties by farmers. Traditionally this was achieved through the following sources of information:

- Secondary sources such as published government reports and available data sets.

- Seed sales and seed multiplication and distribution data.
- Expert opinion and key informant interviews.
- Community level surveys including focus group discussions.
- Farmer elicitation of name and type of variety cultivated as part of household survey.
- Use of morphological descriptors for variety identification.

However, these methods have inherent uncertainty levels and often estimates have wide confidence intervals. DNA fingerprinting offers an alternative approach in which samples from farmers' fields can be matched to known improved varieties and landraces. Use of DNA method is less subjective and it is therefore able to provide a more accurate and credible estimates of adoption rates and associated economic analysis.

2.2 Overview of the DNA fingerprinting workflow

The DNA fingerprinting component of the NIBAS involved an establishment of a clear workflow. The workflow consisted;

1. Establishment of a reference library comprising improved varieties obtained from IITA breeders, National gene bank collection and seed companies.

2. Field sample collection and preservation

3. High throughput DNA extraction that enables extraction of large number of samples per day

4. Sequencing-based genotyping

5. Bioinformatics and cultivar identification
   The report details the analysis and results of the DNA fingerprinting components of the NIBAS project.

2.3 Main activities in the DNA fingerprinting component

The DNA fingerprinting component of NIBAS involved several activities from field sample collection to variety identification analysis. Below are detail reports on the completed activities.

2.3.1 Sample and sample associated data collection

The study sampled households in six states representing the largest cowpea production areas in Nigeria. Leaf samples were collected and preserved in plastic tubes containing silica gel from all the farmers identified varieties growing in each household, provided the varieties were growing in separate fields. In each farmer's field Leaf samples were collected from 12 plants along the longest diagonal to give a total of 12 plants per variety. These leaf samples were bulked and transported to Bioscience laboratory at IITA in Ibadan, Nigeria for DNA extraction. Household information including region ID, state ID, local government area ID, enumeration area ID, household ID and household head's name were captured in a tablet using the SurveyB software. In addition, information on variety name, cropping pattern (mono-cropped or intercropped), field and plot identification, plot size in all the fields owned by the household, the GPS coordinates of the household where survey took place, and the farmer's field were measured.

2.3.2 Establishment of sample tracking system

A standard tracking system is important particularly when dealing with a large sample size to reduce any possible introduction of human errors of sample mismatch and mix ups. Multiple layers of tracking system-using barcode labeled sampling kits and a tablet computer for capturing sample in duplicates and sample-associated information were implemented. This process has improved the accuracy and reliability of the data. Once received in the lab samples, the barcode label on each tube was captured with barcode reader and the samples arranged in a set of 96 on an in-house made plate and assigned distinct plate number. The barcode reader information was crosschecked for any possible error introduced.

2.3.3 Summary on the samples collected

Field sample collection included leaf tissue collection, preservation in silica gel/desiccant containing plastic tubes and recording of sample-associated information. Due consideration and intensive training prior to field visit and on the field were given to enumerators to ensure proper sampling, conservation of plant tissues and capturing sample related information. Before field visit plastic tubes of 250ml size containing 50g of silica gel were prepared and adhesive label with a barcode label unique for each field were prepared in duplicates, one pasted on sample collection tube and the other placed inside the tube. Other sample-associated information was also captured.

2.3.4 DNA extraction and genotyping

DNA was isolated following DNA extraction protocol (Dellaporta *et al*., 1983) with some modification for large-scale sample. All the extracted DNA samples were quantified using spectrophotometer and agarose gel electrophoresis for quality and quantity assessments. All extracted samples that pass the minimum quantity requirement (300ng/µl) were shipped to Diversity array technologies for DART genotyping. Genotyping and SNP calling was done using the proprietary DART sequencing-based genotyping.

2.3.5 Development of a reference library for variety identification

Estimation of adoption rates and intensities through DNA fingerprinting requires a well curated and comprehensive reference library of known improved varieties. In addition, collections representing known farmer-preferred landraces could be included in the reference library to identify common or popular unimproved varieties. In the absence of a useful library, initial DNA fingerprinting can serve to establish a baseline data. We assembled a total of 206 accessions to establish the cowpea reference library. Of these, 27 came from National Centre for Genetic Resources and Biotechnology (NACGRAB) and five seed companies that contributed between one to two varieties each. The remaining 179 accessions were obtained from the IITA cowpea breeding program[1]. Seed companies that contributed cowpea reference materials include Greenspore Seeds, Soy Agricultural Company, Tecni Seeds and Value Seeds. To extract DNA, 12-24 seeds from each variety were planted in a screen house and leaf samples from 12 plants per variety were bulked and DNA extracted at the IITA Bioscience Facility in Ibadan. To assess the pairwise similarity threshold for identification of same varieties, the reference library accessions were redundant genotypes between 2 and 8 times as technical replicates. 136 accessions genotyped twice and 58 genotyped four times[2]. The map of Nigeria showing sampling geolocations undertaken by the NIBAS surveys is shown in Figure 3.

[1] Our attempts to get more references from other national partners were not successful
[2] The reference library used for variety identification contains genotypes from the NIBAS and TLIII projects

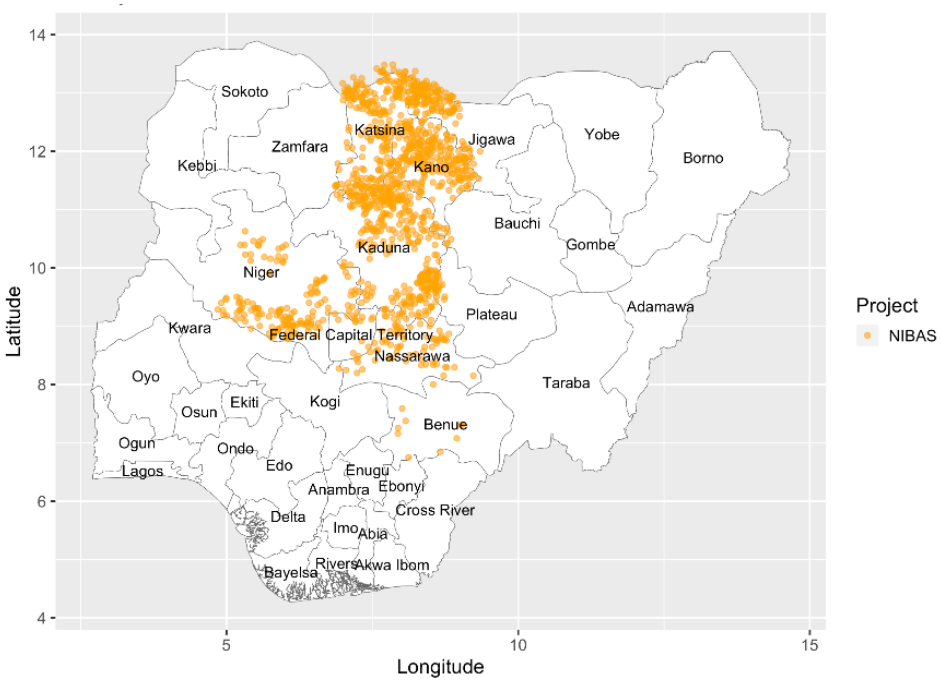

Figure 3: Map of Nigeria showing sampling geolocations undertaken by the NIBAS surveys

2.3.6 Bioinformatics

We received genotyping results for 1757 samples. Reference library made up the rest of the samples which included multiple genotyping of same samples as technical replicates. The total genotyping rate is 96% indication low proportion of overall missing data. The marker data were filtered to remove 7969 non-informative SNPs with minor allele frequency of less than 1%. 103 samples that had more than 40% missing data were removed from further analysis.

2.3.7 Cluster analysis and variety identification

Genetically identical sets of accessions were identified from pairwise genetic distance and hierarchical clustering methods implemented. Field samples are subsequently matched to corresponding known varieties in the reference library. Pairwise identity-by-state distance was calculated using 2208 SNPs and 1757 samples that passed the quality control filters. An ad hoc distance threshold for determining accessions of same cowpea varieties was empirically established using data from technical replicates involving the reference library accessions. Individuals in the library were genotyped between two and four times to establish the residual distance due to technical SNP errors. 136 accessions were genotyped twice and 58 genotyped

four times. A frequency distribution of pairwise distance revealed that the appropriate ad hoc distance threshold for calling two samples as identical was 0.1 (Figure 4)[3].

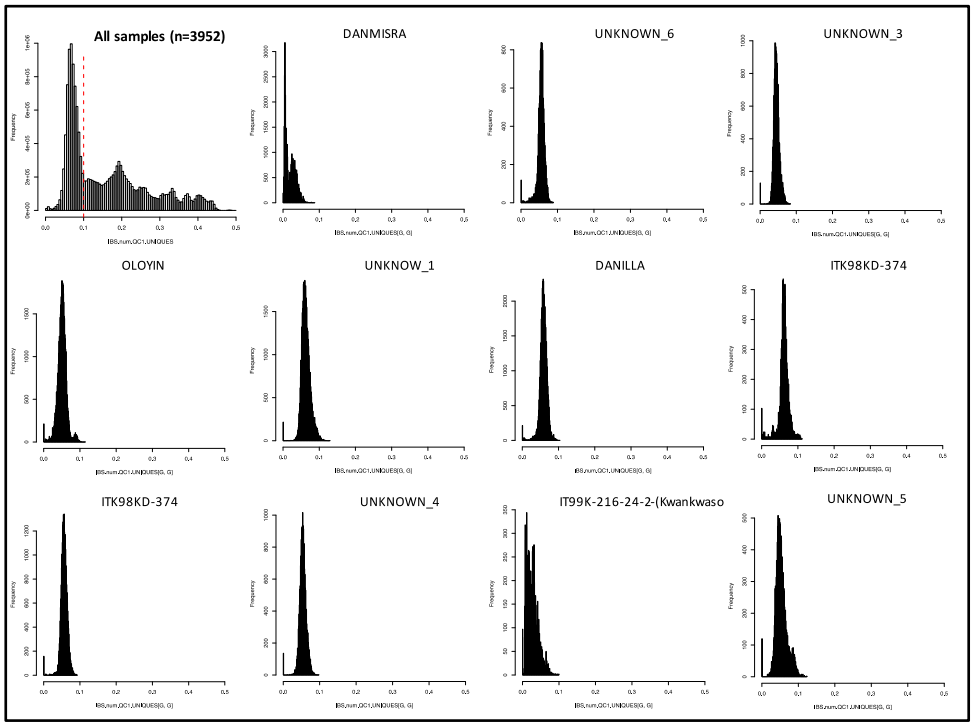

Figure 4: Histogram of pairwise distance within groups of accessions of the same variety as identified through DNA analysis.

Based on this cut-off similarity, we were able to cluster samples into unique sets of genetically unique cultivars. The corresponding threshold in the hierarchical clustering dendrogram (Ward's method) was 0.65 (Figure 5). We chose the Ward's clustering method since it minimizes the total within-cluster variance and is therefore able to group accessions that are genetically similar more appropriately. The same thresholding criteria was applied to the combined reference and field sample data to identify the varieties cultivated by farmers.

[3] Note that, only varieties with more than 100 individuals represented are shown in the figure. Note that the maximum pairwise distance for all samples go up to 0.5 while the distance within the main sets of identified varieties do not exceed 0.1.

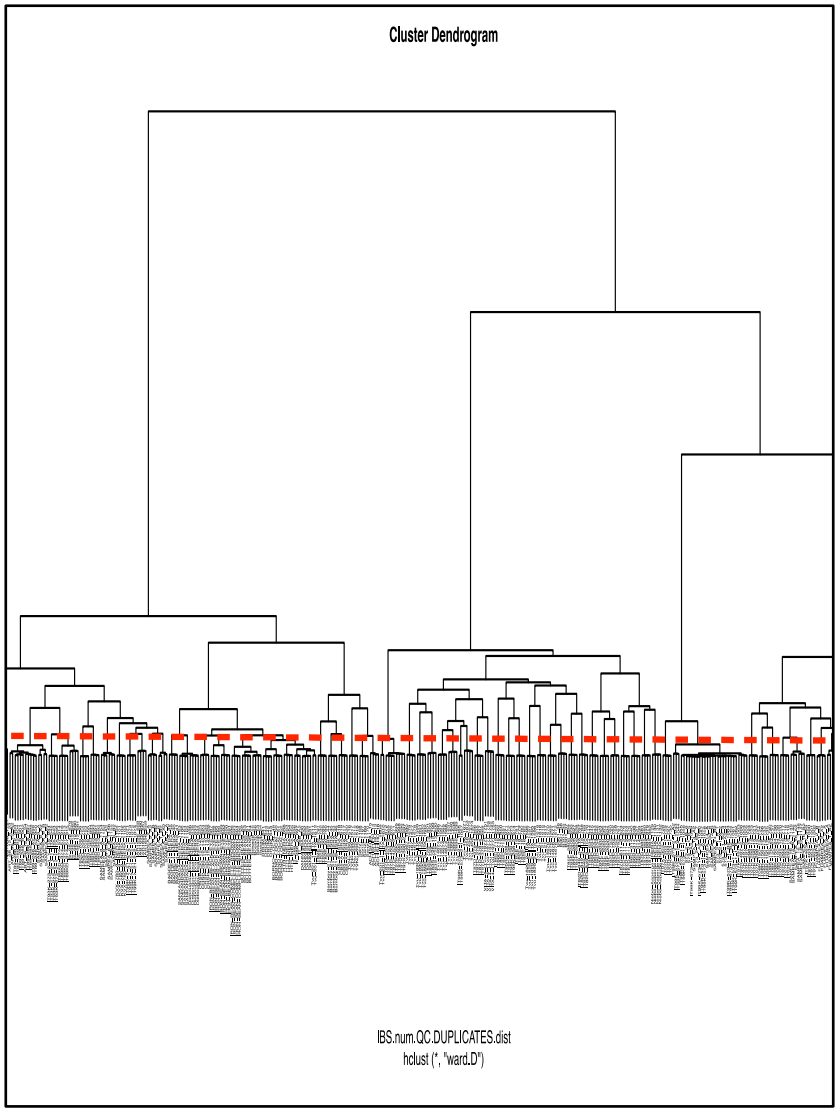

Figure 5: Established criteria for determining identical sets of clones based on 2 samples genotyped in duplicates

Individual samples from NIBAS surveys did not form independent clusters which would have indicated differences in SNP content despite being processed and genotyped independently. In addition, there was no noticeable differences in the genotype profiles from pooled and individually genotyped accessions of the same variety. This shows high levels of genetic fixation/purity in cowpea and conforms to the expectation of self-pollinated species. Hierarchical clustering revealed two major groupings (Figure 6)[4]. The first is composed of mainly accessions that matched unimproved/unknown cultivars in the reference library. The second major cluster

[4] Note: (A) accessions highlighted in red are from NIBAS. (B) Field samples in grey and library samples in yellow.

consists mainly of improved varieties (samples with prefix IT). Accessions were considered same variety if they clustered together (Ward's dendrogram height < 0.1 and IBS distance < 0.1). These thresholds correspond to those determined from technical replicates of the reference library. Using the *ad hoc* pairwise distance threshold of 0.1 for variety identification, we uncovered unique 55 varieties.

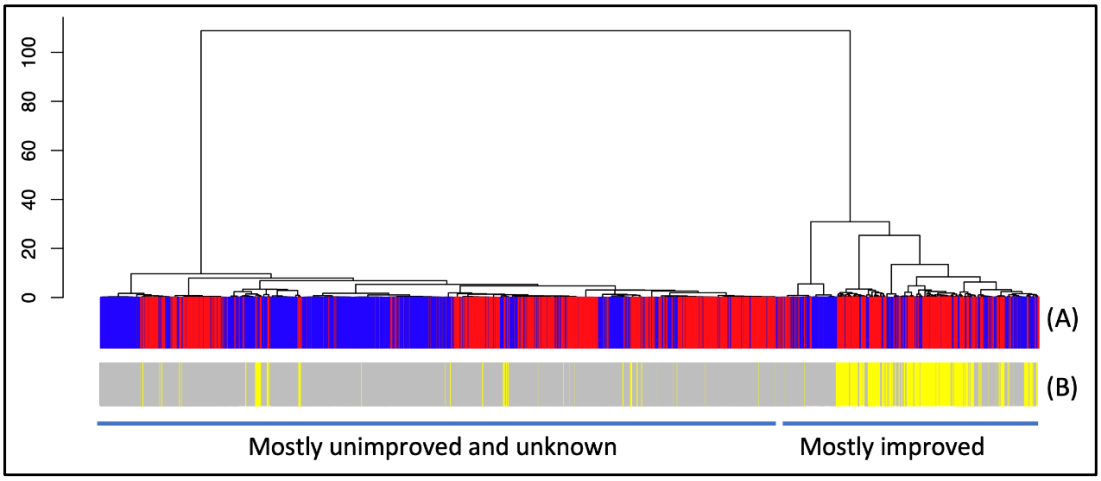

Figure 6: Hierarchical cluster dendrogram of the NIBAS samples.

2.3.8 Identified varieties

The field samples were subsequently identified based on the reference library accessions that co-occurred in the same cluster (Table 2).

Table 2: Summary of improved varieties

Nr	Matched variety	# of samples	Status
1	IT89KD-374	68	Improved
2	IT99K-216-24-2	67	Improved
3	IAR353	52	Improved
4	SAMPEA_2	54	Improved
5	IT98K-131-2	28	Improved
6	IT89KD-288	13	Improved
7	IT81D-985	25	Improved
8	IT97K-1042-3/IT97K-1042-8	19	Improved
9	IT99K_573-2-1	11	Improved
10	IT90K-277-2	13	Improved
11	IT95K-222-3	6	Improved
12	IT99K-573-1-1	2	Improved
13	IT84S-2163	7	Improved
14	IT97K-49935	2	Improved
15	IT97K-1069-6/IT97K-1069-8	1	Improved
16	IT07K-292-10		Improved
17	FUAMPEA_1	2	Improved
18	IAR355	3	Improved
19	Ife-Brown	2	Improved
Total		375	

Of the total 1757 field samples from NIBAS, 652 could be matched to a known variety in the reference library and 375 of these were improved varieties while 277 were categorized as unimproved. The remaining 1105 were not matched to any accession in the reference library and are referred to as "Unknown". Of the identified cultivars, 18 matched known improved varieties in the reference library (Table 2). The most common improved variety uncovered in the study was IT89KD-374. The second most common improved variety was IT99K-216-24-2. Other identified improved varieties were IAR353, SAMPEA_2, IT98K-131-2, IT89KD-288 and IT81D-985. The remaining 12 varieties occurred with less frequency. Besides these, we were able to match 1013 field accessions to 10 different unimproved varieties in the study (Table 3). The most common were Danilla; DANMISRA and OLOYIN also known as Kannanado Brown.

Table 3: Varieties matching unimproved germplasm in reference library

Nr	Matched variety	# of samples	Status
20	Danilla	26	Not improved
21	DANMISRA	49	Not improved
22	OLOYIN/SILVER-BROWN/KANNANADO-BROWN	45	Not improved
23	IRON-BEAN-COMPLETE-BROWN	76	Not improved
24	IRON-BEAN-BROWN-EYE	32	Not improved
25	Mai-babanda/IRON-BEAN-BLACK-EYE	4	Not improved
26	KVX-309-6G + Suvita2/TN5-78/Tvu_201_1_IITA	32	Not improved
27	Bosadp/BUTTER-BEANS	7	Not improved
28	DRUM/BORNO-LOCAL	1	Not improved
29	Kannanado	5	Not improved
Total		277	

2.3.9 Seed purity

Cowpea is primarily self pollinating species with relatively low levels of within individual heterozygosity and within variety genetic diversity. Lack of genetic purity due to rare outcrossing events or seed mixing could compromise DNA-based variety identification. Seed purity was assessed by calculating individuals accession homozygosity i.e. number of homozygous SNPs divided by total number of genotypes SNPs per individual (Figure 7). The average homozygosity across all samples was 0.90 (standard deviation = 0.07) and falls within the expected range for self-pollinated species. Relatively low homozygosity level in our samples results from pooling DNA from 12 individuals from each field plot into a single sample for genotyping. We chose this approach because it reduces the cost of genotyping by a factor equal to the number of individuals in the pooled sample. By genotyping pooled DNA, we were able to

detect diversity within plot most likely due to seed mixture. However, this did not reduce our ability to cluster accessions based on overall pairwise genetic similarity.

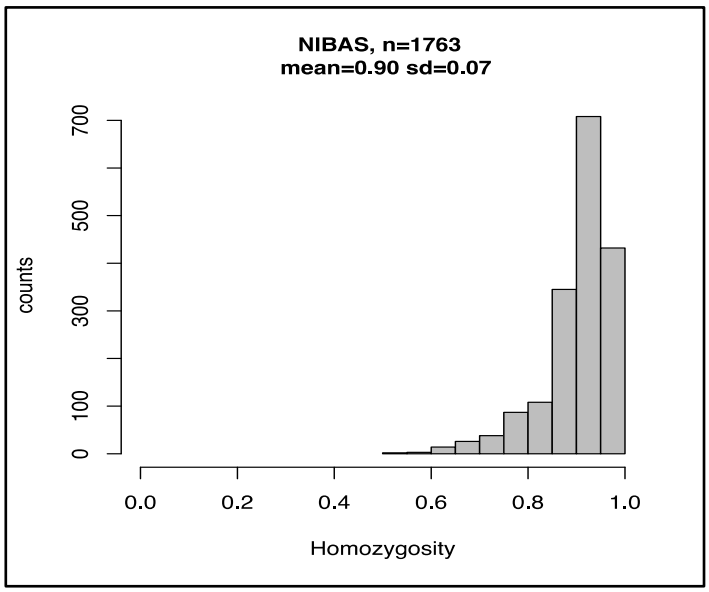

Figure 7: Histogram of within-individual homozygosity

3. Comparison of adoption rates based on matched self-reported and DNA data

3.1 Adoption rate: Binary measure

This section presents adoption rate of improved cowpea varieties using the matched self-reported and DNA-fingerprinted adoption data. Adoption is defined based on the proportion of households growing at least one improved variety. That is, a farmer is considered as an adopter if he/she grows at least one improved cowpea variety in one or more of his/her plots. In addition to the binary adoption measure, we also reported intensification rate, the proportion of total cowpea area under improved cowpea varieties, using both self-reported and DNA-fingerprinted adoption data. Results are reported in Table 4 for the whole sample. The result shows close to 23.2% adoption rate for the whole sample (i.e., about 23% of households grow at least one improved variety in one or more plots) when using farmers' self-reported data.

For the DNA-fingerprinting, categorizing varieties into improved and unimproved was not straightforward. Specifically, we encountered groups of varieties in the field samples that could not be matched to any reference sample. In this report, we preferred to categorize these as unmatched/unknown rather than declare them as unimproved. The large proportion of the total sample is classified as unknown/unidentified due to absence of matching varieties in the reference library. This lack of match could be attributed to incompleteness of the library. The

lack of complete identification of field samples is thus more likely due to the incompleteness of the library rather than being landraces. Maintaining viable seed copies of every released varieties is not trivial particularly for seed crops that require regular seed regeneration. It is likely that many original improved varieties were lost through natural attrition over the years from the collections of the various breeding programs. Alternatively, it is also likely that many of the unknowns are truly unimproved varieties based on inference from the hierarchical clustering analysis since they cluster together with known local varieties such as Danmisra, Oloyin and Danilla (Figure 5). As such, some of the varieties in these group could be improved and hence we considered the adoption rate based on the matched and identified varieties, as a lower bound. With the above caution in mind, adoption rate as measured by DNA-based varietal identification is about 24.4%.

Table 4. Adoption rate of improved cowpea varieties full sample (%).

	Proportion of households (%)		
	Adopter	Non-adopter	
Survey	23.2	76.8	
Proportion of households (%)			
	Adopter	Non-adopter	Unidentified/unknown
DNA	24.4	16.4	67.4

In Table 5, we present adoption rates for each of the six states separately. Adoption rate seem considerably low across all states. In the second column (DNA), results are based on the assumption that all the unidentified/unknown varieties are unimproved.

Table 5. Adoption rate of improved cowpea varieties (%).

	Self-reported Adoption rate (%)	DNA Adoption rate (%)
Nasarawa	23.8	27.7
Benue	36.6	35.8
Kaduna	15.4	10.8
Niger	20.4	27.2
Kano	42.2	36.3
Katsina	11.6	21.6

3.2 Adoption rate: Area under improved varieties

In this section, we reported the intensity of adoption, which measures the total cowpea area under improved cowpea varieties, in Table 6. Based on self-reported data, the intensification rate for the whole sample is about 16%. However, the intensification rate stands at 17% based on

DNA-fingerprinting, implying at least 17% of the cowpea area in the sampled area is under improved varieties (i.e., assuming that all the cowpea area under unknown/unidentified varieties is unimproved).

Table 6: Intensification rate (proportion of area under improved cowpea varieties)

	Survey intensification rate (%)	**DNA intensification rate (%)**
Full sample	16.3	17.4
Nasarawa	28.6	20.3
Benue	32.5	31.9
Kaduna	12.6	8.5
Niger	14.3	13.9
Kano	30.4	28.6
Katsina	9.5	18

3.3 Varietal turnover

Analysis of the survey data did not uncover single dominant cowpea variety with the most common genotype (IT89KD-374) accounting for only 5% of the total cowpea area in the sample. Cumulatively, however, the identified improved varieties make up only 17.2% of the total cowpea area in the sample (Table 7).

Following the approach of Brannen and Byerlee (1991), we also calculated the area weighted average age (AWA) of the identified improved varieties listed in Table 1 as follows:

$$AWA = \sum_{i=1}^{N} T_i \, p_i \qquad 1.$$

where T is the time period (in years) since variety i was released (i.e., calculated as 2018 - release year), p is the proportion of area allocated to crop variety i. Note that, since the release date of the unidentified varieties is unknown, we excluded these groups of varieties in our calculation. In the NIBAS sample, the area weighted average age of improved varieties is only 20.2 years. This result suggests that, on average, it takes about 20 years for a given variety to be replaced by another variety.

In estimating AWA, we rely on two key assumptions: First, we assumed that the age (i.e., T) of improved varieties that were not officially released is zero. Thus, we re-estimate WTA with the assumption that these varieties were 30, 20 or 10 years old and found that the corresponding AWA would have been 29.9, 26.6 and 23.4 years, respectively. Second, we assumed that all the unidentified varieties are unimproved and hence were excluded in the AWA estimation. To

provide a broader perspective, we re-estimate AWA with the assumption that these varieties are improved and released 30, 20 or 10 years ago[5]. In this case, the AWA corresponding to the assumption of 30, 20 or 10 years of age becomes 29.6, 20 and 12.1 years, respectively. The AWA estimates were sensitive to the assumed ages of unknown varieties as these varieties account almost 65% of the total cowpea area in the sample.

Table 7: Adoption rate and age of most popular cowpea varieties

Identified Variety	Release year	Type	Area in the sample (ha)	Share (%)	Cumulative (%)
IT89KD-374	1991	Improved	58.27	4.98	4.98
IT99K-216-24-2	Not released	Improved	34.34	2.94	7.92
IAR353	1979	Improved	25.38	2.17	10.09
SAMPEA_2	1979	Improved	23.49	2.01	12.10
IT98K-131-2	Not released	Improved	10.42	0.89	12.99
IT89KD-288	2009	Improved	6.78	0.58	13.57
IT97K-1042-8/ FARV-13	1971	Improved	7.53	0.64	14.21
IT81D-985	Not released	Improved	14.08	1.20	15.41
IT98K-491-4/IT90K-277-2	2005	Improved	3.33	0.28	15.70
IT99K_573-2-1	2011	Improved	3.64	0.31	16.01
IT84S-2163	Not released	Improved	2.88	0.25	16.26
IT95K-222-3	Not released	Improved	2.32	0.20	16.45
IAR48/SAMPEA_5/IAR355	1979	Improved	1.03	0.09	16.54
FUAMPEA_1	2016	Improved	4.13	0.35	16.90
IT97K-497-2/SAMPEA_15	2011	Improved	0.18	0.02	16.91
Ife-Brown	1970	Improved	0.97	0.08	17.00
IT04K-321-2/	Not released	Improved	0.93	0.08	17.07
SAMPEA_10	2008	Improved	1.49	0.13	17.20
Danilla	-	Local	9.51	0.81	18.01
Anmisra	-	Local	30.83	2.64	20.65
Oloyin/Silver-Brown	-	Local	30.71	2.63	23.28
Iron-Bean-complete brown	-	Local	68.54	5.86	29.14
Iron-Bean-Brown-eye	-	Local	13.41	1.15	30.28
Iron-Bean-Black-eye	-	Local	2.88	0.25	30.53
KVX-309-6G + Suvita2/TN5-78	-	Local	37.92	3.24	33.77
Bosadp/Butter-Beans	-	Local	5.62	0.48	34.25
DRUM/BORNO-local	-	Local	0.02	0.00	34.25
Kannanado	-	Local	2.85	0.24	34.50
Unknown/Unidentified	-	-	766.15	65.50	100.00
Total			1169.7	100	

[5] In addition, we assume the different varieties in the unknown group represent the same variety.

3.4 Misclassification rates

Next, we report variety identification by farmers by comparing responses from the survey with the results of the DNA-fingerprinting analysis. Based on our matched self-reported and DNA fingerprinted data, we calculated the rate of misclassification at the household level. These include:

i. Reporting/misperceiving local varieties as improved (i.e., False positives),
ii. Reporting/misperceiving improved varieties as local (i.e., False negatives)

As shown in Table 8, only 6.3% of the farmers correctly identified the improvement status of the cowpea varieties they planted. For instance, about 15.8% farmers misidentified improved varieties as local (i.e., false negatives). These group of farmers "*believe that they are growing local varieties but actually grow improved varieties."* Similarly, about 2.6% farmers misidentified local varieties as improved (i.e., false positives). These group of farmers "*believe that they are growing improved varieties but actually grow local varieties."*

Table 8: Misclassification rate of adoption status at the plot level.

		HH surveys	
DNA		Adopter (%)	Non-adopter (%)
	Adopter (%)	6.3	15.8
	Non-adopter (%)	2.6	12.4
	Unidentified/unknown (%)	12.4	50.5

Further, the DNA-fingerprinting result was inconclusive in identifying the varieties reported as improved by 12.4% of the households. That is, the varieties that were reported as improved by 12.4% of the households during the survey were categorized as unidentified/unknown as they could not be identified by DNA-fingerprinting analysis. Similarly, the varieties that were reported as local by 50.5% of the households during the survey were categorized as unidentified/unknown as they could not be identified by DNA-fingerprinting analysis. The above results also raise several questions. In particular, why do farmers misreport improved varieties? This would require additional information about the cowpea seed system in the sample area. Information such as: Do farmers always purchase seeds for improved cowpea varieties? Do farmers rely on informal seed exchange? Do farmers recycle own seed from year to year? In this section, we examine patterns of misclassification focusing on the type of seed used by farmers (i.e., (i.e., purchased, recycled, obtained from other farmers, etc.). In Table 9, we thus reported improvement status of the identified varieties by self-reported seed source.

Table 9: Misclassification by seed type and sources

Based on self-reported seed source				
Reported seed type	No. Obs.	Improved (%)	Unidentified (%)	Unimproved (%)
Own\recycled	887	20.6	63.4	15.9
Other farmers	41	24.4	56	19.5
Local market/shops	147	27.2	63.3	9.5

In Table 10, we examined the correlation between seed sources and misclassification using a Probit model. The dependent variable (i.e., *Correct improved*) takes a value of one if the variety is reported to be improved by DNA-fingerprinting and zero if the variety is unknown/unidentified. The results reported in Table 10 suggest no correlation between seed sources and variety identification.

Table 10: Correlation between misclassification and seed sources/types

	Own\Recycled
Correct improved	-0.033
	(0.028)
No. observation	1149

Notes: Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients are marginal effects

In Table 11, we present the correlation between genetic quality (i.e., improvement status) and productivity using an unconditional regression. Results reported in Table 11 suggests a positive correlation between adoption of improved cowpea varieties and productivity.

Table 11: productivity and misclassification (unconditional regression results)

	OLS estimates
Improved	0.127**
	(0.061)
Note: local varieties are the reference category	
No. observation	1827

Notes: Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.5 Geographical distribution of the identified varieties

The geographic distribution of the major varieties identified in the survey are presented in Figure 8.

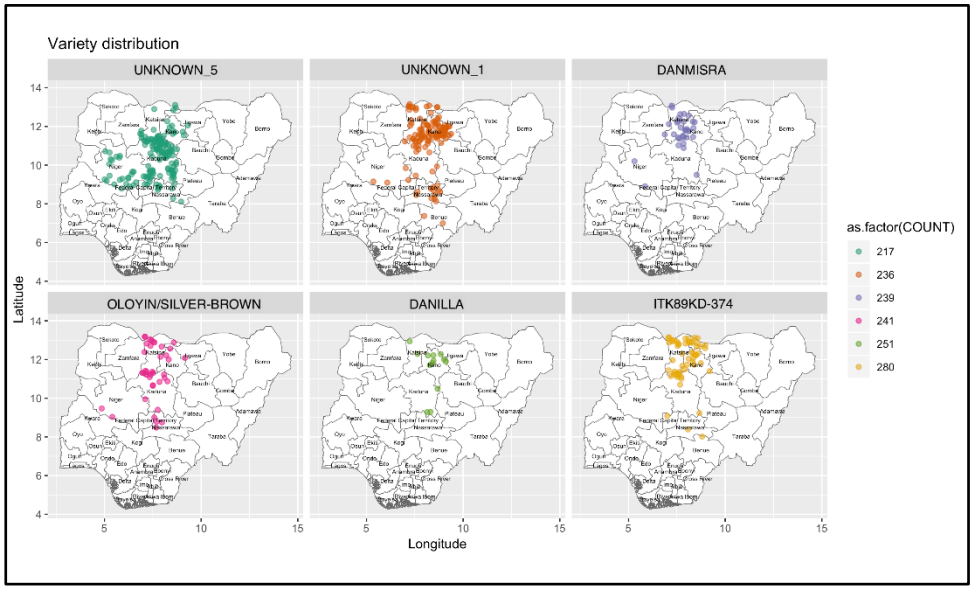

Figure 8: Distribution of the major cowpea varieties

Majority of the common cowpea varieties were found to occur across a large part of Northern Nigeria, from Kebbi in the East to Gombe and Borno in the East. This suggests that these varieties are adapted to these regions that share common agroecological characteristics. Some varieties particularly, the less common genotypes, were found in smaller geographical area.

4. Discussion and recommendations

This report documented the adoption rate of improved cowpea varieties using survey and DNA-fingerprinting based approaches. Based on DNA-fingerprinting analysis, about 24% of the farmers in the NIBAS sample are growing improved cowpea varieties. The adoption rate based on self-reported adoption status from the survey data is about 23%. However, categorizing varieties into improved and unimproved in the DNA-fingerprinted sample was not straightforward. In fact, about 67% of the genotypes collected from the farmers field could not be matched to any of the samples in the reference library. This lack of match could be attributed to incompleteness of the library. As such, these varieties were categorized as "unmatched/unknown". To maximize benefits, farmers need to continuously replace such varieties with seeds from new and superior varieties that have higher genetic potential.

Otherwise, high adoption rates of unknown varieties may not translate to higher productivity. We argue that agricultural development programs and policies should address not only the quantity of adoption (low adoption rates) but also the quality of adoption. This requires efforts beyond mere dissemination of improved varieties to achieve productivity growth and calls for quality control and dissemination of highly quality seeds as an integral part of the dissemination process. Further analysis to understand the nature of the "unknown/unidentified" varieties is also important to disseminate newer and better improved varieties in the areas where these varieties are popular.

5. References

Brennan, J.P. and Byerlee, D. 1991. The rate of crop varietal replacement on farms: measures and empirical results from farms. *Plant Varieties and Seeds*, 4:99-106.

Dellaporta, S.L., J. Wood, and J.B. Hicks. 1983. A plant DNA mini-preparation Version II. *Plant Molecular Biology Reporter,* 1:19–21.

6. Annex

6.1 Varietal turnover using the pooled NIBAS and TLIII sample

In this section, we report an estimate of average age of cowpea varieties using the pooled TLIII and NIBAS sample. Like the NIBAS project, the TLIII project collected adoption data using self-reported and DNA-fingerprinting approaches from 10 States, Borno, Bauchi, Gombe, Jigawa, Kaduna, Kano, Katsina, Kebbi, Sokoto, and Zamfara, which represent about 75% of the total cowpea production in Nigeria. Table 12 shows the list of popular cowpea varieties, year of release and area under each variety. Based on the pool data, about 21.2% of the total cowpea area in the sample is under improved varieties while about 41.6% of the total cowpea area in the sample is under unknown/unidentified varieties. The rest, about 37.2% is still under local varieties. The area weighted average age (AWA) of the identified improved varieties in the pooled sample is about 20.12 years.

Table 12: List of most popular cowpea varieties in the pooled NIBAS and TLIII sample

Identified Variety	Release year	Type	Area in the sample (ha)	Share (%)	Cumulative (%)
IT89KD-374	1991	Improved	404.7	12.34	12.34
IT99K-216-24-2	Not released	Improved	116.3	3.54	15.88
IAR353	1979	Improved	33.98	1.04	16.92
SAMPEA_2	1979	Improved	23.5	0.72	17.63
IT98K-131-2	Not released	Improved	24.62	0.75	18.38
IT89KD-288	2009	Improved	17.75	0.54	18.92
IT97K-1042-8/ FARV-13	1971	Improved	7.53	0.23	19.15
IT81D-985	Not released	Improved	14.1	0.43	19.58
IT98K-491-4/IT90K-277-2	2005	Improved	3.5	0.11	19.69
IT99K_573-2-1	2011	Improved	10	0.31	19.99
IT84S-2163	Not released	Improved	3.5	0.11	20.10
IT95K-222-3	Not released	Improved	7.3	0.22	20.32
IAR48/SAMPEA_5/IAR355	1979	Improved	1	0.03	20.35
FUAMPEA_1	2016	Improved	4.8	0.15	20.50
IT97K-497-2/SAMPEA_15	2011	Improved	7.66	0.23	20.73
Ife-Brown	1970	Improved	0.98	0.03	20.76
IT04K-321-2/	Not released	Improved	6	0.18	20.95
SAMPEA_10	2008	Improved	2.2	0.07	21.01
SAMPEA_16	2015	Improved	5.1	0.16	21.17
Danilla	-	Local	315.7	9.62	30.79
Anmisra	-	Local	328	10.00	40.79
Oloyin/Silver-Brown	-	Local	256.2	7.81	48.60
Iron-Bean-complete brown	-	Local	87.8	2.67	51.27
Iron-Bean-Brown-eye	-	Local	53.9	1.64	52.91
Iron-Bean-Black-eye	-	Local	48.9	1.49	54.40
KVX-309-6G + Suvita2/TN5-78	-	Local	40.3	1.23	55.63
Bosadp/Butter-Beans	-	Local	25.14	0.77	56.40
DRUM/BORNO-local	-	Local	44.6	1.36	57.76
Kannanado	-	Local	20.4	0.62	58.38
Unknown/Unidentified	-	-	1365.52	41.62	100.00
Total			3280.8	100	