



RESEARCH  
PROGRAM ON  
Livestock

*More meat, milk and eggs by and for the poor*

# Progress Narrative: Platform for African Dairy Genetic Gain

Okeyo Mwai, International Livestock Research Institute (ILRI)

*Note: This research report is an extract from the ACGG BMGF report covering the period 1<sup>st</sup> December 2016 – 30<sup>th</sup> November 2017*





Use this form to provide updates to your foundation program officer regarding progress made toward achieving your project's stated outputs and outcomes.

The Progress Narrative must be submitted in Word, as PDFs will not be accepted.

### General Information

<b>Investment Title</b>	Platform for African Dairy Genetic Gain		
<b>Grantee/Vendor</b>	International Livestock Research Institute (ILRI)		
<b>Primary Contact</b>	Okeyo Mwai	<b>Investment Start Date</b>	October 26, 2015
<b>Feedback Contact<sup>1</sup></b>	Okeyo Mwai	<b>Investment End Date</b>	October 30, 2018
<b>Feedback Email<sup>1</sup></b>	o.mwai@cgiar.org	<b>Reporting Period Start Date</b>	December 01, 2016
<b>Program Officer</b>	Donald Nkrumah	<b>Reporting Period End Date</b>	November 30, 2017
<b>Program Coordinator</b>	Elizabeth Weaver	<b>Reporting Due Date</b>	March 31, 2018
<b>Investment Total</b>	\$9,130,404	<b>Opportunity/Contract ID</b>	OPP1130995
<b>Scheduled Payment Amount (If applicable)</b>	\$2,871,274		

<sup>1</sup> Feedback Contact/Email: The full name and email of the contact whom foundation staff queries for various surveys.

### Submission Information

By submitting this report, I declare that I am authorized to certify, on behalf of the grantee or vendor identified on page 1, that I have examined the following statements and related attachments, and that to the best of my knowledge, they are true, correct and complete. I hereby also confirm that the grantee or vendor identified on page 1 has complied with all of the terms and conditions of the Grant Agreement or Contract for Services, as applicable, including but not limited to the clauses contained therein regarding Use of Funds, Anti-Terrorism, Subgrants and Subcontracts, and Regulated Activities.

<b>Date Submitted</b>	16-05-2018	<b>Submitted by Contact Name</b>	Okeyo A. Mwai
		<b>Submitted by Contact Title</b>	Project Director/Principal Investigator
		<b>Submitted by Contact Email</b>	o.mwai@cgiar.org
		<b>Submitted by Contact Phone</b>	+254 733 562795

## Progress and Results

### Progress Details

Provide information regarding the current period's progress toward achieving the investment outputs and outcomes as well as the work planned or anticipated for the next period. In addition, submit the Results Tracker with actual results as requested.

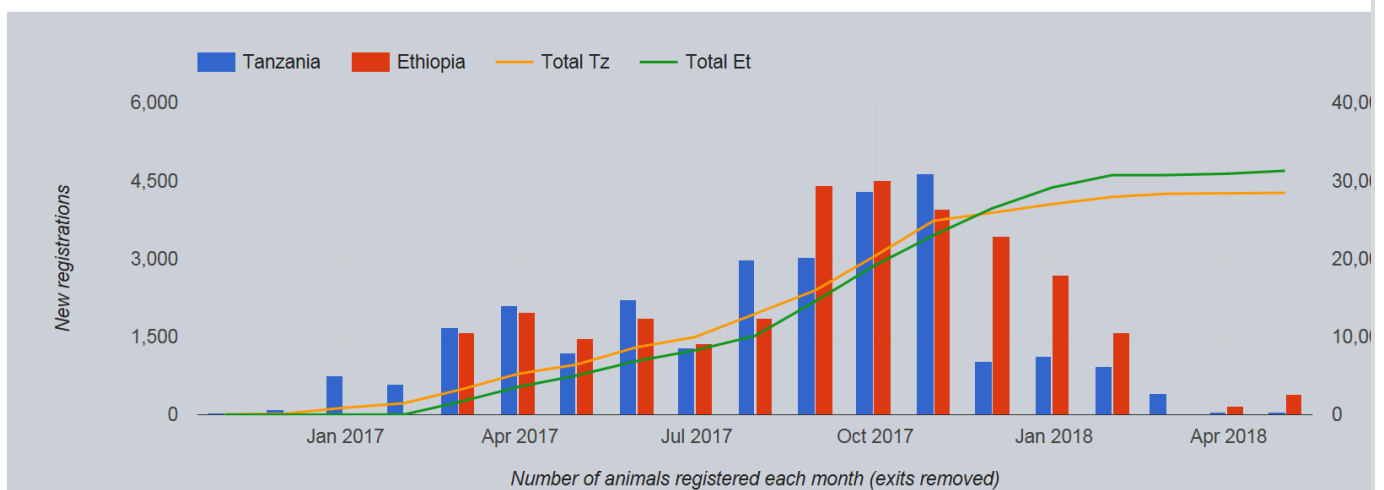
The project has three objectives: 1) to establish National Dairy Performance Recording Centers (NDPRCs) in each of the two project countries; 2) to develop identification, certification systems for crossbred bulls & promote their use for AI and natural service and 3) to establish on-farm information and communication technology (ICT) and digital platforms for capturing data and delivering dairy management information to farmers & service providers. Below find summaries of the progress that has been made on each of the above objectives during year 2 of the program implementation:

1. Establishment of National Dairy Recording Centers (DPRCs) in Ethiopia and Tanzania and that by the end of this reporting period, it was expected that, in each country a DPRC that had been established in Year 1 is operational and is being used to collect data on cows kept by the small holders in the 2 countries. In addition, that the farmers through the DPRCs are receiving feedbacks and farmer education messages. As planned, the 2 DPRCs (one each) are functioning and with National Artificial Insemination Center (NAIC) hosting Ethiopia's DPRC. In **Annex 1** are maps showing the locations and numbers of smallholder herds which are registered on the DPRCs. Due to the current increased visibility and its anticipated and future role, NAIC has since been elevated into a National Livestock breeding and Improvement Institute, with expanded roles and mandates.

In both Tanzania and Ethiopia young scientists have been identified and trained in data handling and analyses, as part of the sustainability plans (*see details under section 3.1.4 below*). In addition, the program has continued to engage the relevant departments (Agricultural and Livestock bureaus), being key stakeholders, in Ethiopia and Local Government Councils (LGCs) in Tanzania and agreed on the respective roles and contributions to the running of the DPRCs. For example, in Ethiopia, all the ADGG's Performance Recording Agents (PRAs) and the sites coordinators who supervise the PRAs are members of staff for the regional Governments working for the Agricultural and Livestock bureaus, while in Tanzania all the PRAs and site coordinators work for the LGCs and are seconded to ADGG.

The project continues to work very closely with the ministries responsible for livestock development, the national research systems, and the collaborating institutions, specifically the Land O' Lakes, and the local partner company the Animal Breeding East Africa (ABEA) (Tanzania) for which the platform provides the recording and monitoring of the performance of the franchised artificial insemination technicians (AITechs). In response to the special needs of the AITechs, most of whom had difficulties in handling more sophisticated digital data capture forms, New simplified PAID and ADGG forms (**Annex 1a, 1b and 1c**) have been developed and successfully rolled out for field use. The report by the Industrial and Scientific Advisory Committee on the program is provided as **Annex 9**.

In Ethiopia, agreement have been reached with LUKE on the merger of the old LUKE-NAIC data and the new digitally collected data has been building up (*see graphs below*), following training of the country partners on the data capture tools.



In both countries, the following additional related activities have been successfully undertaken: a) installation of new equipment and additional (new) performance recording agents (PRAs) and 1 ICT expert in each country have been recruited to support data capture and feedback systems. So far, more than 43,000 herds; 66,000 animals and more than 75,000 inseminations undertaken by PAID's AITechs have been captured/monitored on the two DPRCS ([http://azizi.ilri.org/cgi-bin/ADGG\\_dash6.py](http://azizi.ilri.org/cgi-bin/ADGG_dash6.py); <http://metabase.data.ilri.org/>). Thousands more herds would have been enrolled on the platforms, if the intermittent political unrest

in Ethiopia had not persisted, which curtailed free movements of the PRAs and the AITechs in the field. In Tanzania, 9,690 test-day milk records is now available (**Annex 2a**), with the number of test day record per animals ranging between 1 and 10 (**Annex 2b, and 2c**). Good progress is being made and these records, together with the pedigree records and the genomic relationship information that will be obtained from the 7,000 animals sampled are sufficient to allow estimation of genomic breeding values with reasonable levels of accuracies. It should be noted that new herd registrations had to be halted momentarily, except registrations coming in via insemination records so that the PRAs could focus on collecting in-depth herd/cow performance data and feedback sessions. We therefore expect more data than those presented in **Annex 2a,2b and 2c**. In Ethiopia, about 33,000 records have been registered on the new platform, in addition to the 15,000 cows that the NAIC-LUKE manual system registered. In total, about 174,000 milk records are now available in the Ethiopia's platform awaiting the genomic information later in the year to generate genomic breeding values and raking of cows and bulls on this basis for the first time in Ethiopia.

Farmers have so far received more than 2 million education text messages to help them improve their dairy husbandry practices, and most of them have been satisfied with these some indicating that some herd productivity improvement are already being realized. More farmers want to join-in, and in response, more PRAs have been recruited. In Ethiopia 25 new PRAs have now been employed and equipped with new motorbikes to use for herd recording and providing feedbacks to move things forward quicker given that the political situation has improved. Digital data capture tools have since been revised. In response to feedback by the field users, the initial data capture ODK tools that were developed in year 1 have since been refined through joint efforts by the GDT and ILRI programmers. The following **new digital tools** have been developed or refined and rolled out:

**Tools for Tanzania:**

2. New combined form for PAID Tanzania available in English and in Kiswahili: <http://data.ilri.org/collect/adggtanz/forms/PAID-TZN-Service-SWA-ENG-20180308>
3. New ADGG Tool for Registration of data collectors, AI techs, and all Site coordinators for Tanzania: <http://data.ilri.org/collect/adggtanz/forms/ADGG-REG01-20180305>
4. New ADGG Tool for Monitoring animal and farm performance: <http://data.ilri.org/collect/adggtanz/forms/ADGG-LNG-20180305>
5. New ADGG tool for Registration of Farmers, their farm characteristics and all animal categories: <http://data.ilri.org/collect/adggtanz/forms/ADGG-REG02-20180305>

**Tools for Ethiopia:**

6. New combined form for PAID Ethiopia: <http://data.ilri.org/collect/adggeth/forms/PAID-ETH-Service-00-20180116>
7. New ADGG Tool for Registration of data collectors, AI techs, and all Site coordinators for Ethiopia: <http://data.ilri.org/collect/adggeth/forms/ADGG-ETH-REG01-20180424>
8. New ADGG Tool for Monitoring animal and farm performance: <http://data.ilri.org/collect/adggeth/forms/ADGG-ETH-LNG-20180424>
9. New ADGG tool for Registration of Farmers, their farm characteristics and all animal categories: <http://data.ilri.org/collect/adggeth/forms/ADGG-ETH-REG02-20180424>

In addition to the above, tools for data capture on large scale farms: MS-Access data capture tools were developed which lift historical data to the platforms. Training was provided to partners on the use of the tools, and these are being used by large scale farmers to collate animal performance and pedigree information on their farms, which are then uploaded directly to the ADGG server. Also a tool for collecting hair samples and which integrates the information with the related performance/pedigree data, where the latter exist, and to the ILRI's Azizi <http://azizi.ilri.cgiar.org/> Bio-repository data system, from where the generated data would be access by a wider global scientific community, especially those from the Center for Tropical Livestock Genetics and Health (CTLGH) program. A tool for communication and monitoring progress on the ground, Trello, is currently being used by both the PAID and ADGG data-teams for communication with PRA's and AI technicians in both countries.

2. Establishing on-farm information and communication technology (ICT) and digital platforms for capturing data and delivering dairy management information to farmers & service providers: Farmer information and management support services reaching thousands of farmers in Tanzania and Ethiopia.

2.1. GDT's iCow or other service tests improve service, assesses impact, and adopts best elements for long-term delivery:

2.1.1. All the planned activities, except the integration with Ethiopia's Agricultural Transformation Agency (ATA), have been successfully accomplished. These include:

10. The Farmers' needs were identified and routinely monitored through the following: a) analyses of baseline data collected from the farmers during participatory workshops with farmers and using a theory of change protocol (**Annex 3a**). The reports on theory of change workshop interactions with key stakeholders (**Annex 3b**); b) monthly interactions of the performance recording agents with the farmers; c) through quarterly sites and regional stakeholder platform convenings, and d) through direct text messages received directly to the iCow platform from farmers. These processes helped inform customization of messages and automation of feedback loops.
11.
  - a. Integration of iCow platform with Tanzanian Dairy Performance Recording center (DPRC)
  - b. Developed and signed memorandum of Understanding (MoU) with the Tanzanian host institution-TALIRI.
    - o Hosting of the iCow-TZ section in the cloud
  - develop USSD menu codes and SMS codes for iCow Tanzania and Ethiopia
  - Customize and translate iCow contents into Kiswahili and Amharic for Tanzania and Ethiopia respectively
  - Integrate iCow calendar onto iCow ADGG Tanzania and Ethiopia
  - Automate feedback loop with optimum efficiency, by linking farmers directly back the iCow platform for more appropriate and relevant actions.
  - Demonstration of the value of feedbacks through farmer adoption and object assessment of productivity improvement is on-going through a design MSc Thesis study and comparison of baseline performance compared to the average year 2 herd performance levels. However participating farmers are reporting improved husbandry practices.

To test potential uses of automated phenomics under local smallholder systems, Ice Robotics in Scotland in partnership with SRUC, ILRI and the University of Nairobi initiated the testing of accelerometer technology to determine the proportion of time an animal is active by measuring lying and standing time and of subsequent bouts of each of these activities. It also generates the step count within the given period. The activity levels are calculated using a proprietary "Motion Index" and an alert given when a cow is detected on heat. Results are relayed in an easy to ready dashboard for decision making and correlated milk sampling for milk constituent analyses is also undertaken. Informative dashboards <https://cowalert.com/heats/>. A brief outline of the design is given in **Annex 4a**. The results are being collated and will be reported in project's year 3 report.

### 3.1. Analytical tools to process data to generate information for farmers feedback are in operations

A report by the SRUC partner on the development of lactation curves using the DGEA data is given in **Annex 4b**. The analytics tools for farmer feedback mostly on benchmarking were ready for testing by middle of April 2017. Initial test in old database indicated scripts were working, but given the scanty data points, results showed that predominantly, the lactation curves from the smallholder farms were flat (**Annex 4c**). With more data points being generated from the new longitudinal ADGG and NAIC-Ethiopia-Luke, better lactation curves will be developed using the same algorithms to develop more robust Key Performance Indicators (KPI), and results used thereafter to routinely support bench-marking and farmers' feedback and alert systems.

#### Other new tools and products developed by the ADGG Program team:

1. To overcome the perennial lack of internet connections in Ethiopia, **SMS module for ODK tool** that enables data collected on ODK to be transmitted to the DPRC servers without using Internet connection has been developed and tested. This would be used in areas with low or no internet connection, but its full deployment awaits further wider discussion and approval by the relevant Ethiopian authorities.
2. **A Telegram-based chatbot that allows farmers and PRA's to interact directly** with the DPRC has been developed and is being tested in Kenya and Tanzania. It currently provides instant feedback on milk test-day data for any household or location aggregated over any time period. It also provides hyper-localized weather forecasts through collaboration with aWhere.com and allows PRA's to see nearby households on their tablets. The intention is to expand to provide farmers and PRAs a real-time snapshot of our engagement with a given household or locale, which we believe, will encourage engagement. Through a separately-funded (CGIAR BigData platform) initiative farm.ink are developing tools to provide additional farm management advice and information through a chatbot interface and when this is ready it will

be combined with the ADGG chatbot. The chatbot can be discovered at @ILRI\_ADGG\_bot. The need for a data connection is rapidly declining in importance as a limitation, at a minimum, PRA's can pre-prepare reports on their tablets to discuss with farmers..

3. **A market place product** (iCow Soko): <http://icow.co.ke/tags/soko> that connects dairy cow sellers and buyers, and which runs on USSD platform has been developed and is fully operational in Kenya, and is being populated with data for testing ahead of its operationalization in Tanzania. The platform would be then linked to the breed composition and genetic ranking tools, catalogues and platforms, so that buyers can ascertain the breed compositions and value of cows and heifers being offered for sale hence potentially creating an income stream for the DPRCs.

We consider that rapid and transparent access to project information is key to ensuring engagement. Our strategic use of a mix of SMS and data based communication tools provides great flexibility and scalability. It provides PRAs with realtime access to project data, and positions us to exploit the rapidly increasing availability of smartphones to farmers while not neglecting farmers without access to data.

### 3.1.1 Models for genetic and genomic evaluation

Models for genomic predictions for young bulls using GBLUP are already in place and tested giving accuracies that varied from 0.4 to 0.30. Data from ADGG sites are continuously being examined and similar models will be run when the data builds-up to a useful amount, especially when the new pedigree and genomic information is received.

New models fitting dominance and fitting breed proportion as separate effects using a multi-trait approach have been developed and tested using DGEA data. Because of the limited dataset, the dominance effects were essentially zero, but this is expected to rise when additional ADGG data becomes available and analyzed.

The multi-trait approach resulted in a slight improvement in the predictive ability of the model, although the accuracy of prediction did not improve. From the SNP-BLUP models, the proportionate contribution to the GEBV top cows with high exotic genes and top cows with high indigenous genes were examined. The highest contributions were from chromosomes 14, 8, 1 and 2 for top cows with high exotic genes while these were chromosome 3,1,10 and 5 for top cows with high indigenous genes. This points out that using the appropriate analyses, separation of chromosomal segment of a crossbred animal's haplotypes into either ancestral parental breed (i.e. exotic dairy and indigenous one) is possible, but the haplotype length must be long enough. This is important in further refining what linked traits are likely to be contained in a crossbred bulls and cows in addition to the other attributes such animals may have. This is also highlights the need to improve the multi-trait approach by fitting different SNPs which originate from either the exotic and indigenous breeds.

The UNE team are undertaking within-chromosome admixture analysis to allocate different genomic regions to exotic or indigenous origin. These will be used in a GWAS study and hence locate the specific genomic regions associated with milk production in the exotic and indigenous breeds.

Using the previously collected related data the analytical models are being developed, tested and perfected as the data density increases, in readiness for finalization upon receipt and inclusion of the genomic relationships data. The ultimate results of which will be the genomic ranking of the bulls and finally selection and recruitment of the top best bulls into AI centers.

### 3.1.2 Reduced chip for genomic prediction.

This work has been led by the UNE team, but with inputs from the ILRI scientists. Initially an efficient method for SNP selection was developed based on the covariance among SNPs weighted by their minor allele frequency (MAF). This gave better results compared to other selection methods such as evenly spaced markers or maximization of MAF or random selection. This was followed by examining various imputation software and the impact of several factors on imputation. Results of our study (**See Annex 5**)- (Aliloo et al., 2018, titled "The feasibility of using low density marker panels for genotypes imputation and genomic prediction of crossbred dairy cattle of East Africa"- Manuscript JDS-18-14621; Accepted J.D. Sci. 2018) confirms that the genotypes of East African crossbred dairy cattle can be imputed with sufficiently high accuracy to achieve highly accurate GEBV, thus providing large savings in genotyping costs, or increased effectiveness of selection for fixed genotyping budget.

There is no great advantage from including purebred reference breeds in the crossbred imputation but, if they were available, use of purebred animals that are the actual ancestors of the crossbred animals might provide a significant increase in accuracy. Imputation methods for SNP selection which are population-based were more efficient when cross-bred animals and pure breeds were included in the reference population and when genomic relation between reference and target population was high. A set of about 3-4k SNPs are adequate for imputation and genomic prediction. This will be verified, later this year, when the low-density



chip for breed composition and parent verification which is currently being developed, is field tested in Kenya, later this year. A report on the first levels of stakeholder engagement is contained in **Annex 10**.

### 3.1.3 Sampling and genotyping

Currently, hair samples have been collected from more than 6,000 cattle (females and males) in Tanzania. These include about 5,000 cows with lactation data in the ADGG and about 1,000 active males in the project sites. The hair samples have been shipped to Neogene for genotyping. A similar exercise is in progress in Ethiopia, this was hindered by the restrictions on movements due to the then prevailing political situation in the country, but situation in the country is now calm and sampling is on-going. The Ethiopian samples together with the those from selected Kenyan crossbred cows/calves and active bull populations will be submitted for analysis in May 2018.

### 3.1.4 Capacity building of DRPC

Training on genetic evaluation and data handling was help last year at Ethiopia for partners and other personnel associated with the DRPC in each country. This was implemented in collaboration with the colleagues from LUKE in Finland. Practical session used example data from the Luke data base (**see Annex 6**). The next training is scheduled for June 2018. The annual reports being prepared and submitted by the partner countries continue to improve, as shown by the Tanzanian report (**Annex 7**).