



Better lives through livestock

Feasibility of joint genomic evaluations for smallholder dairy data in Tanzania and Ethiopia

R. Mrode , C.C. Ekine-Dzivenu, J.M. Ojango, A.M. Okeyo

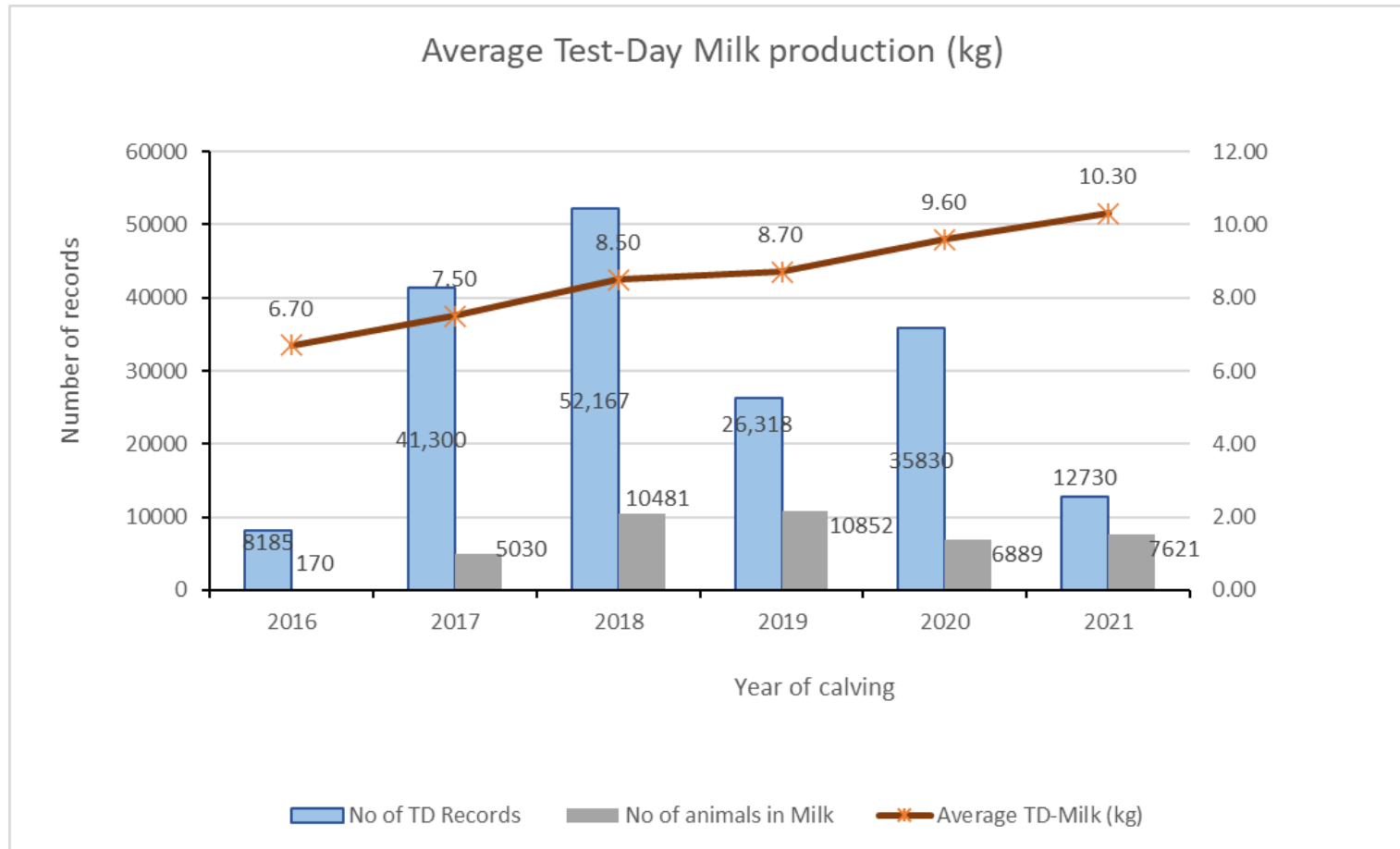
Interbull Opening meeting, Montreal , Canada, 30-31 May
2022

Background information

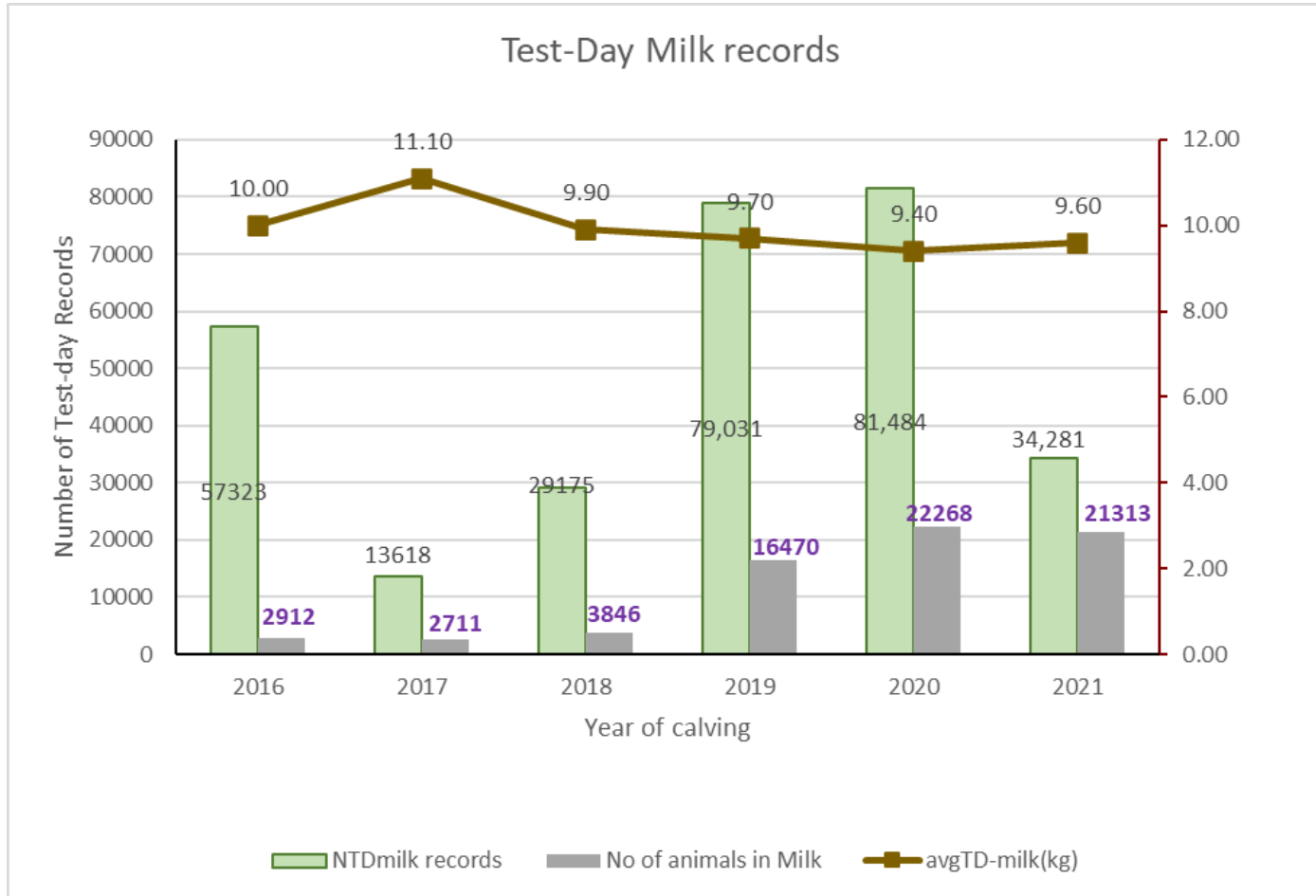
- African Dairy Genetic Gain (ADGG) project commenced in 2016
- Tanzania and Ethiopia, extending to Uganda, Kenya and Rwanda in the next phase
- **Tanzania**
 - Commenced routine data collection and process
 - Routine genomic prediction initially GBLUP and ssGBLUP
- Ethiopia
 - Built in a previous project by LUKE, Finland- data and pedigree available
 - Expanded data collection and implement genomic prediction on ssGBLUP



Summary data on main countries in African Dairy Genetics Project - Tanzania



Summary data for main countries in African Dairy Genetics Project - Ethiopia



Genotyping Strategy

- GeneSeek Genomic Profiler (GGP) Bovine 50K used for genotyping:
- About 5600 animals genotyped in both countries using hair samples
- After QA, 40581 SNPs remained, and these were imputed to HD
- Just of about additional 3000 animals genotyped this year

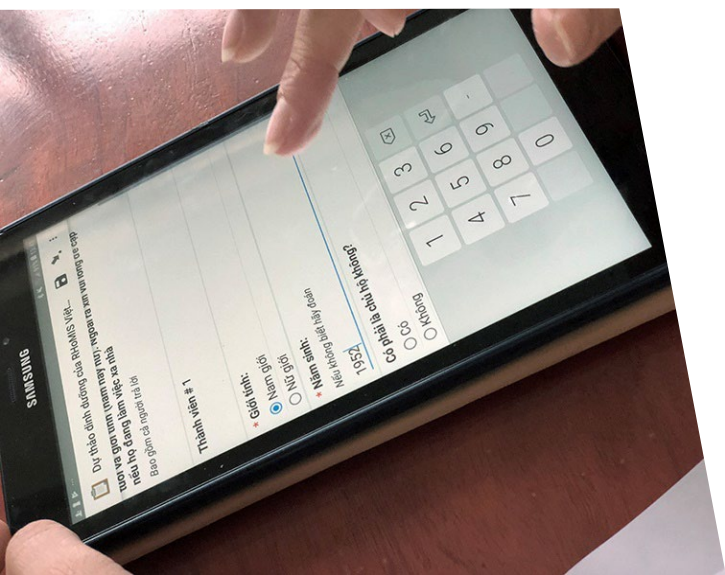


Data structure

- Considering cows at least 3 test days records with genotypes gives us these data sets
 - Tanzania: 1916 cows with 19599 test days records
 - Ethiopia: 1642 cows with 16530 test days records
- With such small data sets, the questions was whether the accuracy of prediction could be increased by combining data from both countries

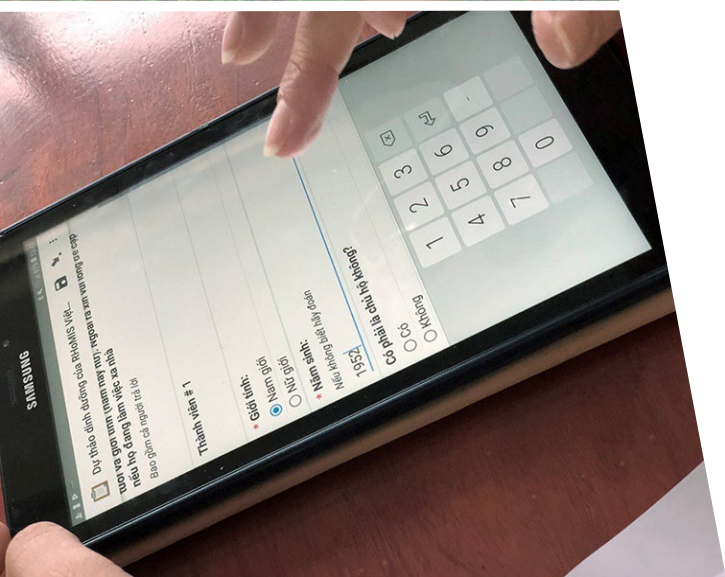


Small reference population



- We know that:
- Pooling data within breeds
 - Effective -- Intergenomics
- Pooling data across breeds: Effectiveness depends on
 - Genetic distance between the breeds
 - Density of marker chip
 - Genomic prediction method – Bayesian mixture models

Small reference population



- Effectiveness have not been examined from cross-breeds populations
- This study examines pooling data from Ethiopia and Tanzania
- Issues:
 - Connectedness uncertain as breeding having driven by government policies and agents of AI companies
 - Lack of pedigree in Tanzania and renumbering of foreign bulls used in Ethiopia
 - Lack of pedigree implies genetic links can only be determined from markers : GBLUP or SNP-BLUP
 - Needed for ADGG next steps as project wants to explore across country evaluations

Within Country Analysis

- **GBLUP - Fixed Regression model** : Fixed ward, age nested parity, test-year-season, fixed curves with Legendre polynomials nested within breed classes by parity plus random herd animal and
- Cows classified to 4 breed classes based on proportion of exotic genes : >0.875 , $0.61-0.875$, $60-36$, and < 0.36
- Validation data set comprised of 276 and 349 cows born after 2014 in Tanzania and Ethiopia respectively
- Accuracy was computed as correlation between YD of validation cows and their GEBV



Analyses of combined data

- Examine level of connectedness
 - PCA analysis
 - Estimate genetic correlation
- YD from the within country analysis were used in all analyses on combined data
 - Bivariate analysis
 - Fitting country as fixed effect and random animal effect
- $\text{Var}(\mathbf{a}) = \mathbf{G} \begin{pmatrix} \sigma_{a11}^2 & \sigma_{a12} \\ \sigma_{21} & \sigma_{a22}^2 \end{pmatrix}$ and $\text{Var}(\mathbf{e}) = \mathbf{R} = \begin{pmatrix} r_{11} & 0 \\ 0 & r_{22} \end{pmatrix}$

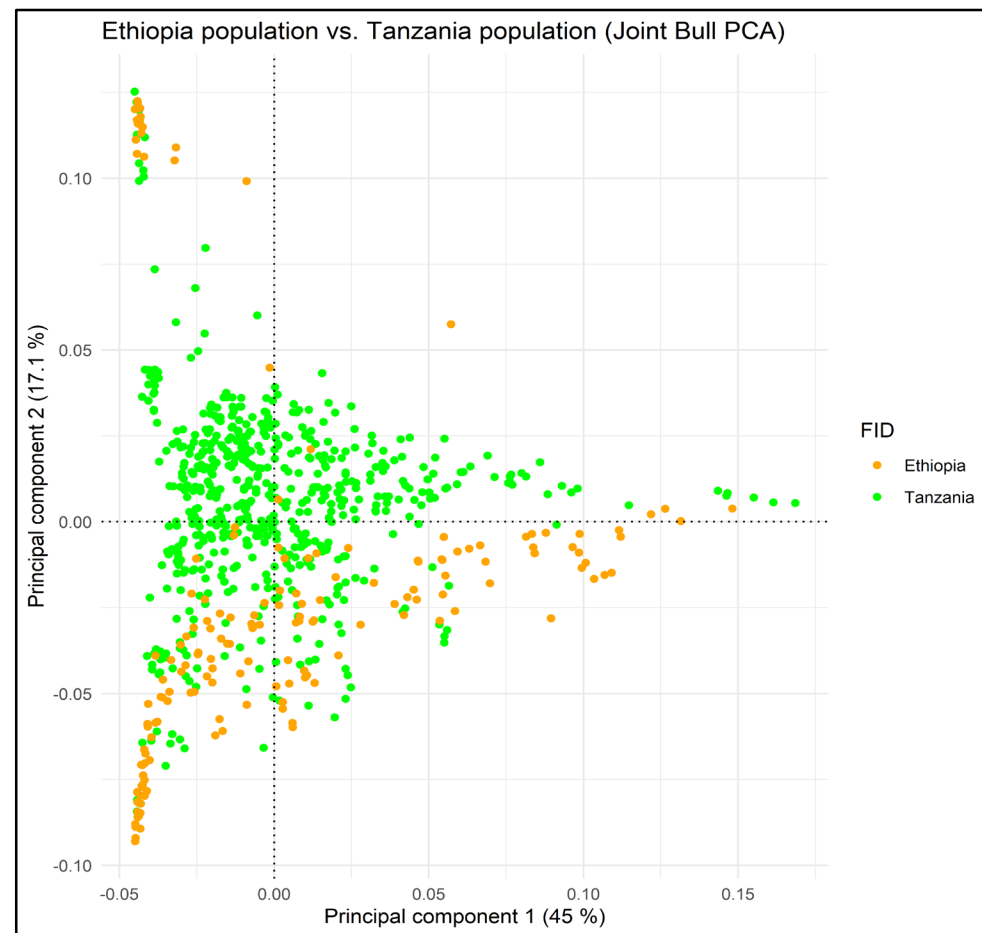
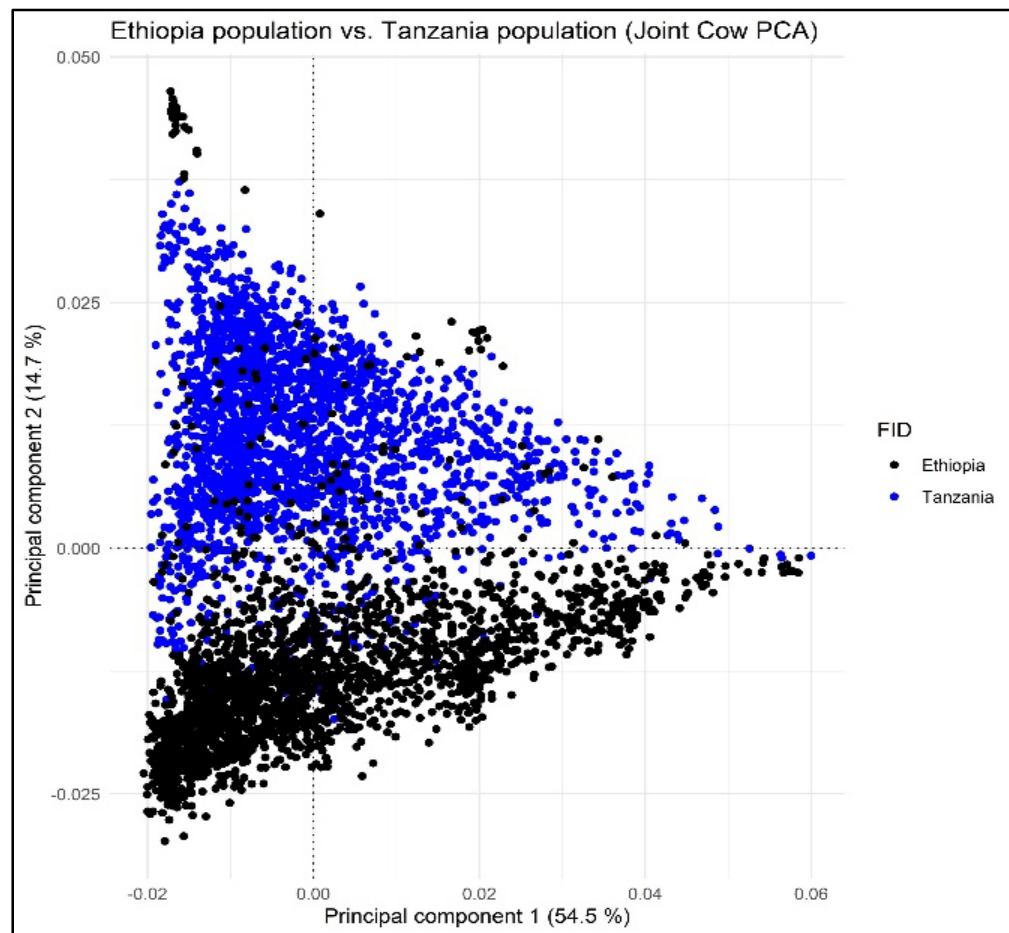


Analysis of combined data

- Assume genetic correlation of unity between both countries
- Combine all data and genotypes
- Sub-setting:
 - Combining only the data and genotypes of the top 100 to 500 Tanzanian animals most related with Ethiopia data to the Ethiopia data and vice versa
 - Degree of relatedness computed from **G** matrix



PCA analyses and genetic correlation



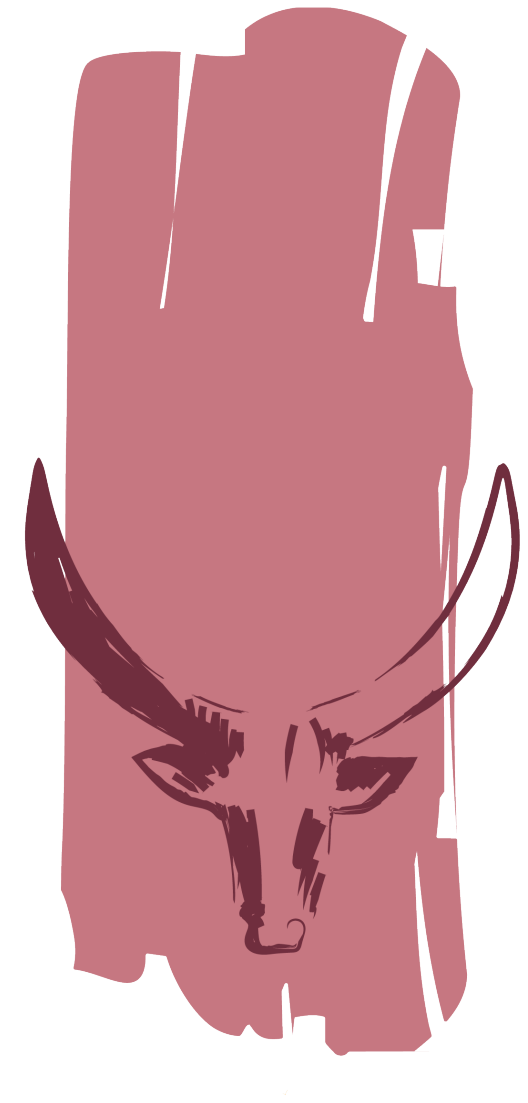
Based on 3558 cows and 745 bulls

Estimate of genetic correlation from bivariate analyses was 0.12 ± 0.10

Validation results for milk yield

Tanzania: $h^2 = 0.11 \pm 0.03$, Ethiopia $h^2 = 0.05 \pm 0.02$

Data set	Tanzania		Ethiopia	
	Regression	Correlation	Regression	Correlation
Within Country Data	1.10	0.58	0.66	0.39
Combined Data	1.17	0.58	0.67	0.40
Sub-setting				
Within Country + 250	1.01	0.59	0.7	0.39
Within Country +500	1.01	0.59	0.7	0.40



Conclusions

- Across country genomic prediction not feasible given the current data structure
- Study was useful in terms of helping ADGG design the next steps
 - Planned exchange of semen from top ranking bulls from ADGG within country evaluations will be necessary
 - Strategically expanding activities in Kenya as most East Africa countries have historically bought animals from Kenya
 - Design of ADGG collaborative projects with two International AI companies planning to test some of their bulls in Africa



Acknowledgements



BILL & MELINDA
GATES *foundation*

**Dairy Farmers & Farmer
organizations**

African Dairy Genetics Gain

Innovative use of genomics and information technology
to improve the productivity and profitability
of the small holder dairy farms in Africa



Thank you