GENEMAP AFRICA BIOBANKING TRANSFORMATION AND FUTURE DIRECTIONS

Oluwafemi G. Oluwole PhD, Chandre Oosterwyk BS, Samuel Adadey Mphil, Khuthala Mnika MS, Abdoulaye Yalcouye MD, Edmond T. Wonkam MD, Noluthando Manyisa MS, Kalinka Popel BS, Mario Jonas MS, Mzodidi Tutuka BA, Vicky Nembaware PhD and Ambroise Wonkam MD, PhD

Department Of Pathology, Division Of Human Genetics, University of Cape Town, Anzio Road, Observatory, Cape Town, South Africa

Background

The consolidation of research collaborations in Africa is critical to reduce the level of global disease burden on the continent and enhance capacity (Chang et al). Improving the knowledge and implementations of Laboratory Information Management System (LIMS) can validate quality data and build research capacity. LIMS is a computer-based application used for managing samples, procedures, equipment, reagents, staff and performing analysis (Skobelev et al), to achieve end-point quality data that can be used for generating commercial reports (Skobelev et al).

GeneMAP Africa is one of the largest indigenous research initiatives that investigates the pathology of sickle cell disease and hearing impairment in the African populations. The initiative has collaborators in about ten different African countries. This study describes the implementation of LIMs across four study sites in Africa to enhance efficiencies and data validation of ~ 10, 000 different biospecimen.

Methods

BAOBAB tool (*https://baobablims.org*) was implemented for the data storage and cataloguing, as described in (Fig. 1). Research personnel from Cameroon, Ghana, Mali and South Africa were trained (Fig. 1) on shipping procedures, data capturing (barcoding system) sampling ID using unique code (Table 1) to curate complex pedigrees (Fig. 2), as well as experiment monitoring and storage of biospecimen. A standard operating procedure that combined pre-analytical annotation, quality stratification and a meta-data platform (BAOBAB, REDCap and the biorepository support of the Clinical Laboratory Services (CLC) (*www.cls.co.za*), was outlined to ensure that biological samples are collected adequately and catalogued according to the ISBER logistics guidelines 2009. The preservation of biological samples at the point of collection adhered to the principles described by the Division of Laboratory Sciences, National Center for Environmental Health, CDC, USA. Then, we performed timeous DNA or RNA extraction using specific kits. The quality of the sample extracted upon delivery were differentiated from stale samples using Nanodrop ™ and gel electrophoresis techniques.

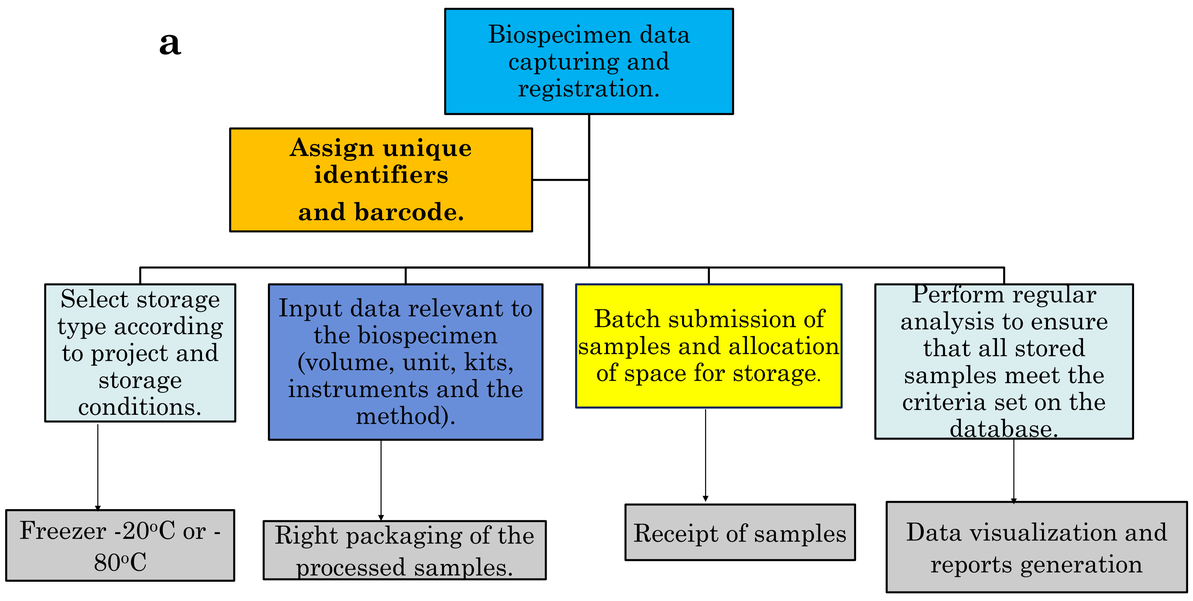

Fig. 1 : (a) LIMS flow chart (b) members of GeneMAP group at a training session

Results

Of the 2000 biospecimen collected and shipped to Cape Town within 12 months period, no damage occur. The comparison of DNA samples extracted timeously with the stale samples, showed notable quality and concentrations changes that impacted the downstream genetic analyses (Fig. 3). The meta-data (Fig. 4) platform when implemented, would be outputting data based on the similarities and unique structures of the components whenever its being accessed.

Table 1: Unique codes used as identifiers in sampling

Pedigree code	Relationship to proband
P1	Proband
O1	Affected offspring
M1	Affected mother sampled
F1	Affected father sampled
E1	Affected grandpa
G1	Affected grandma
W1	Affected wife
H1	Affected husband
B1	Affected brother sampled
S1	Affected sister sampled
C1	Affected cousin sampled
A1	Affected aunty sampled
U1	Affected uncle sampled
N1	Affected niece
K1	Affected nephew

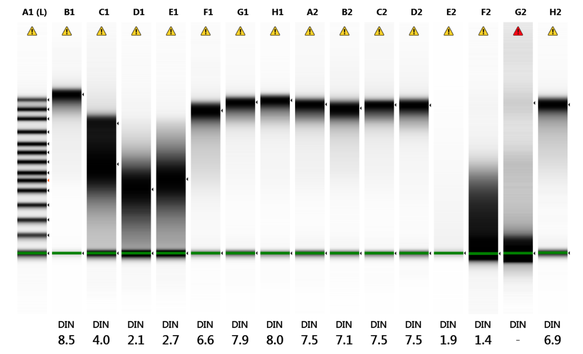

Fig. 3: Gel-electrophoresis of DNA samples done to differentiate quality

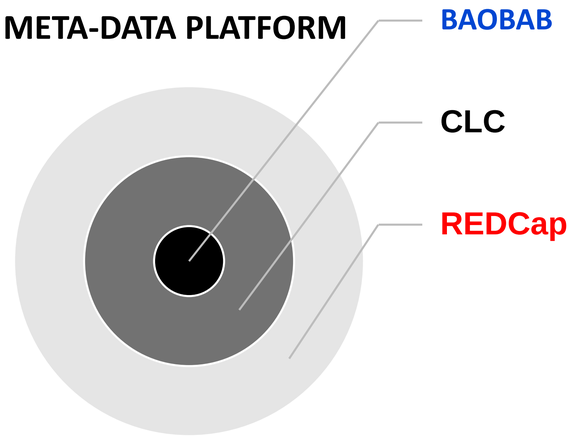

Fig. 4: Synchronized meta-data components

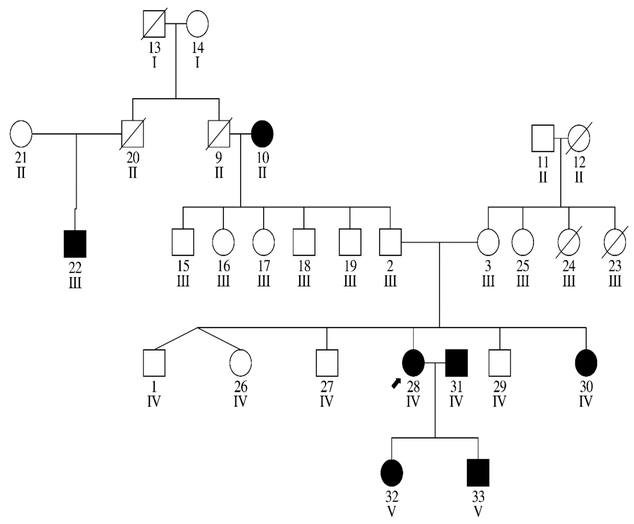

Fig. 2: Annotation of complex pedigree using the sampling method

28IV = GH.NR.CNA.**P1.01.00**
14I = GH.NR.CNA.**G0.01.01**
10II = GH.NR.CNA.**G1.01.02**
2III = GH.NR.CNA.**F0.01.03**
3III = GH.NR.CNA.**M0.01.04**
27IV = GH.NR.CNA.**B0.01.05**
30IV = GH.NR.CNA.**S1.01.06**
31IV = GH.NR.CNA.**H1.01.07**
33V = GH.NR.CNA.**O1.01.08**
25III = GH.NR.CNA.**A0.01.09**

Discussion & Conclusions

The integrity of research is based on adherence to core values Biological samples can be highly susceptible to changes over time and can be easily contaminated if not well stored (National Academic of Sciences). Degraded samples can produce spurious results in research. As we (scientists) are working within a much broader system that profoundly influences the integrity of research results (Skobelev et al), adequate system like the GeneMAP LIMs will be essential in large projects. The interesting thing about our study is that the implementations of the LIMs and SOPs for the validation of research data generated from the biospecimen that we are currently using in the GeneMAP projects have improved drastically. The efficiency rate have improved significantly also, with about 80% accuracy.

References

[1] A. Y. Chang, V. F. Skirbekk, S. Tyrovolas, N. J. Kassebaum, and J. L. Dieleman, "Measuring population ageing: an analysis of the Global Burden of Disease Study 2017," *Lancet Public Health*, vol. 4, no. 3, pp. e159–e167, Mar. 2019.

[2] D. O. Skobelev, T. M. Zaytseva, A. D. Kozlov, V. L. Perepelitsa, and A. S. Makarova, "Laboratory information management systems in the work of the analytic laboratory," *Meas. Tech.*, vol. 53, no. 10, pp. 1182–1189, Jan. 2011.

[3] "Foundations of Integrity in Research: Core Values and Guiding Norms - Fostering Integrity in Research - NCBI Bookshelf," *National Academies of Sciences, Engineering, and Medicine; Policy and Global Affairs; Committee on Science, Engineering, Medicine, and Public Policy; Committee on Responsible Science. Washington (DC).* Available: https://www.ncbi.nlm.nih.gov/books/NBK475948/.