

Bioinformatics Platform: Harnessing Bioinformatics For Food Security And Sustainable Development

Mark Wamalwa¹, Joyce Njuguna¹, Dedan Githae¹, Maureiq Atieno¹, Erik Bongcam-Rudloff², Josephine Birungi¹

¹Biosciences Eastern and Central Africa (BecA) - ILRI Hub, Nairobi, PO Box 30709, 00100, Kenya;

²Swedish University of Agricultural Sciences (SLU), Box 7023, 75007 Uppsala, Sweden;



Bioinformatics

The Science of the Future

Project summary

The BecA-ILRI Hub Bioinformatics platform is a shared resource committed to give access to innovative methods and tools for the genetics and genomics of agronomical crops and animals with specific focus on research projects led by National Agricultural Research Systems (NARS), and universities within eastern and central Africa. It is hosted on a High Performance Computer (HPC) platform located on the ILRI campus. Using this platform, the BecA-ILRI Hub has embarked on research activities including various genome (partial and complete) sequencing projects (genome re-sequencing, viral genomics and metagenomics). Metagenomics projects have been particularly informative, as they provide tools to quickly scan bacterial and viral communities in specific ecosystems with the potential of applications to various food value chains. This robust platform has significantly increased our ability to address major issues related to pathogens discovery, crops improvement and food security and continues to support capacity building and services in the BecA Hub region and beyond.

Outputs

Through the BecA-SLU bioinformatics capacity building partnership initiated in 2012 under the Sweden ministry for foreign affairs (Sida) funding:

- Increased network and data storage capacity to accommodate bioinformatics and life sciences research (Figure 1).
- A standalone Bioinformatics platform (eBioKit) with more than 200 software packages and biological databases (Figure 2) were developed at SLU and transferred to 6 NARs - Cameroon (University of Buea), Kenya (KALRO), Sudan (National Center for Research), Tanzania (MARI) and Uganda (NARO & Makerere).
- Over 100 scientists from NARS were trained in Bioinformatics using the eBioKit through in-country institutional capacity building workshops & group training workshops at the BecA-ILRI Hub (Figure 3).



Figure 1. The HPC platform is a 152 CPU Linux cluster, based on AMD 64-bit Opteron technology with a total of 80TB disk storage.



Figure 2. The eBioKit, a standalone Bioinformatics suite



Figure 3. Researchers from NARs attending a Bioinformatics workshop at the BecA-ILRI Hub.

Outcomes

- A critical mass of African scientists in East and Central Africa (ECA), addressing African agricultural problems using bioinformatics and genomics technologies.
- Access to high-speed computational analyses and big storage capacity due to the revamped state of the art HPC consisting of 152 computing cores, over 400GB of RAM and 80 terabytes (TB) of distributed disk storage.
- The eBioKit suite with more than 200 software packages and biological databases was developed at SLU and transferred to 6 NARs.
- Better support for research programs within BecA and NARs
- Increased number of publications from NARs citing the use of BecA bioinformatics platform in their research. Publications in peer-reviewed journals – Science PMID: 24763584; J. Wildl Dis. PMID: 25380362; PLoS ONE PMID: 25541981; Mycologia 2014 13-186. PMID: 25344263; International Journal of Poultry Science, 13: 279-291; EMBnet.journal 20, e755.

Partnerships

Sida funding of the BecA-SLU bioinformatics capacity building partnership supported various projects highlighted below:

1. Genome sequencing of the Finger Millet (*Eleusine coracana*) and maize stalk borer (*Busseola Fusca*).
2. Pathogen discovery in small farming ecosystems under the plant virome project.
3. Technology transfer to NARs through bioinformatics capacity building workshops.

Key partners include: SLU, ICIPE, Reed College (USA), MARI, IITA and ICRISAT



Potential to scale-up

- The field of genomics is caught in a data avalanche and DNA sequencing has become faster and cheaper at a pace outstripping the rate at which computing gets faster and cheaper (Moore's law).
- There is increasing requirement for compute and storage resources, and the expected need for services for this area of research.
- It is conceivable that the HPC infrastructure and the associated algorithms for analysing molecular data would require 'scaling up' to keep pace with the changing Genomics technologies.
- Our focus is shifting towards development of complex, dynamic and diverse workflows for automating computational tasks.

Mark Wamalwa

m.wamalwa@beca.org

<http://hub.afribiosciences.org/>

Partner institution: Swedish University of Agriculture Sciences (SLU), BecA-ILRI Hub

Funding: Swedish Government

