H3ABioNet Montly Newsletter Issue 6 : October 2013

Foreword

The most significant event for the consortium in October was the annual meeting of the General Assembly, which is comprised of the Principal Investigators. The meeting, held in Johannesburg, was attended by PIs and/or their representatives from almost all the participating institutions. On the first day, nodes reported back on their progress for the year and it was exciting to hear about the impact their involvement in H3ABioNet has made on their teams or institutions. H3ABioNet activities were reported to the Scientific Advisory Board on the second day, and we received excellent feedback from the Board members who were present (Jessica and David). Finally, on the third day of the meeting the working groups met and individuals had a chance to join H3Africa working groups and interact with H3Africa project members. Some of the H3ABioNet PIs stayed on to attend the H3Africa consortium meeting, providing the opportunity for further interaction with project members and for us to gather more information on their bioinformatics needs. The full minutes and presentations from the second annual H3ABioNet meeting can be found on the website at:

http://www.h3abionet.org/events/meetings/8-internal-content/84-2013-h3abionet-annual-meeting-october

Another exciting announcement was the awarding of a supplementary request for funds to build the facility for a data portal to enable H3ABioNet to store H3Africa data prior to submission to the European Genome-phenome Archive. The infrastructure working group are now discussing the best model for data storage. As we move into the second quarter of year 2, we need to focus on moving forward with some of our successes from year 1 and overcoming some of the challenges.

Prof. Nicky Mulder.

H3ABioNet
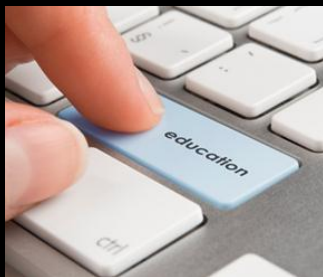
- **Education and Training**

&

The Education and Training working group provided a report to the General Assembly (GA) and Scientific Advisory Body (SAB) on the progress of the group's activities for the first year. A number of topics were also discussed which ranged from:

1) Moving from international trainers to local based trainers
2) Following up with individuals that had been trained to determine:
   a) What topics they would like to becomes trainers in
   b) What additional intervention would be required for them to become experts in the topic they choose to be trainers in
3) Development of a masters style curriculum for bioinformatics
4) Discussion of the mentorship and internship opportunities

H3ABioNet would like to ensure that personnel trained within bioinformatics by the ToT course would form a nucleus of expert trainers that comprise of individuals from North, East, South and West Africa who would be able to provide specialised training and workshops in their topics of interest. To this end, a survey was compiled and sent out to workshop individuals to determine what topics they would like to become experts in and what additional resources they would require to become facilitators within these topics. The SAB was also very enthusiastic about this proposal and suggested to communicate with the trainers in future to determine which individuals are passionate about the subject and enable those individuals to develop their expertise.

The next set of courses will be aimed at students enrolled for a degree programme and a curriculum of relevant topics needs to be agreed upon by the E&TWG group and a Task Force aimed at establishing a bioinformatics curriculum has been established. Some Universities offer dedicated degrees programmes in bioinformatics while the majority of Universities in Africa have students enrolled in other degree programmes but do have a substantial part of their research dedicated to bioinformatics. Hence, by identifying a core set of training topics and providing a suitable curriculum, standardisation can be achieved.

The internship and mentorship programmes were discussed for which the uptake has been rather slow. The difficulty lies in identifying good mentors who are willing to volunteer their expertise although it was decided this can also take the form off a co-supervised projects to collaborations based on specialist subject areas between the various Nodes. The SAB suggested that the application link to the internship programme should be alot more visible.

Dr. Nash Oyekanmi.

- Infrastructure

Infrastructure

The annual activities of the Infrastructure working group were presented at the second annual H3ABioNet meeting to the Node PIs and the SAB. Some of the milestones assigned to the ISWG were determined not to be part of this working group's mandate and in some instances are too broad. The ISWG Chairs have decided to work on the milestones to make them more specific and to attach dates and people to specific milestones. Ownership of specific milestones should belong to a specific working group who will take responsibility while functioning with other working groups.

A lot of the discussion on the specific milestones for Year 2 centred around Galaxy with the intended users and potential workflows to be developed and implemented needing to be identified. Although Galaxy is very useful for non-bioinformaticians, development of workflows and implementation would only be worthwhile if groups and their needs are known as the ISWG might waste a lot of valuable time and resources developing a utility that no one uses. Hence the Galaxy milestones will be examined further by ISWG Chairs and deliberated upon by the ISWG.

The Data management milestones will be important for the ISWG, H3ABioNet and the H3Africa Consortium as H3ABioNet is meant to act as a staging area for the data that will be collected. Hence the actual metadata, storage system and provision of access to the data will have to be managed by H3ABioNet. Additionally, some of the H3Africa PIs might like to have their data presented in a manner that makes mining, visualisation and subsequent analysis easier such as in the form of a BioMart for which a Task Force should be constituted.

The H3Africa data will still take some time before being generated which means that at present the network has some time to develop their skills. The SAB suggested that H3ABioNet should start working with what publically and privately available data the Nodes might have in order to increase the Network's skills and enable development of the various Nodes.

The SAB was particularly interested in the mentioned Iperf / Ping tests that H3ABioNet is seeking to run. At present though, data only exists for a few sites within South Africa and would like to see more sites included in the test with the test occurring repeatedly as the results will be invaluable for IT Directors to motivate to their Institutions for better equipment and infrastructure.

Dr. Alia Benkahla.   Prof. Scott Hazelhurst.

- **Research and Tool Development**

Research and Tool Development

The progress of the Research and Tool Development working group was communicated to the H3ABioNet GA and SAB. A major issue for the RTDWG has been the poor participation in meetings by members which has led to an indepth discussion about the focus of the RTDWG. A number of the initial milestones involved the development of Galaxy workflows and migration to HPC of these workflows for which the RTDWG is not presently equipped to handle. The RTDWG has been strong in starting up and ensuring of collaborations between the various Nodes, as evidenced by the research presentations made to the SAB.

In order for some of the milestones to be met, the RTDWG will work with other groups such as the ISWG who will assume responsibility for some of the milestones. Hence the coming focus of the RTDWG will be more on initiating, encouraging and monitoring collaborative research projects between the Nodes and for the Network as a whole and less on tool development. As the data from the H3Africa projects is not due for some time, the RTDWG will look to the various Nodes to determine which Nodes have data that needs to be analysed and which Nodes have the expertise to provide support / capacity in analysing the data or bringing a research project to fruition.

The RTDWG will also work closely with the E&TWG on encouraging internship and mentorship activities within the Network to support on-going or soon to be initiated collaborative projects. The RTDWG would also like these mentorships, internships and a few workshops to be research oriented so they can result in outputs like conference talks / posters and publications.

A number of research projects such as NetCapDB and SNAP were presented to the SAB for their review with useful feedback being received. A useful comment from the SAB was for the research projects to emphasise how they can be of benefit to the Network and also demonstrate that the long term research activities taking place are a result of the Network and not in response to the Network's deliverables. To that extent, the SAB also suggested that as H3ABioNet is billed as a pan African bioinformatics network, H3ABioNet should work more closely to support projects in Africa that require some bioinformatics expertise both outside and within the Network to promote the idea of a pan African bioinformatics network and ensure its sustainability.

Dr. Julie Makani.   Prof. Ezekiel Adebiyi.

H3ABioNet

- **User Support**

User Support

From the USWG perspective, 2012/2013, the 1st project year, was productive with notable successes, but at times progress got held back by some challenges. We were pleased to report on our milestones with some completed, others already started on but still on-going and a few of them to be reconsidered.

The main success of the USWG presented to the GA and SAB has been the implementation of a fully functioning helpdesk system and the recruitment of individuals to assist with the helpdesk. A lot of discussion took place round the milestones for which it was decided that joint Task Forces between the different working groups should be formed to tackle these milestones but with a single working group being responsible for a specific milestone.

The slow uptake of the helpdesk was discussed with the SAB who provided some recommendations such as making the helpdesk more visible on the home webpage as a resource for bioinformatics support and advertising it widely to the H3Africa consortium and during the workshops with a "Did you know.." type strategy. Suggestions were made by the SAB to not restrict the helpdesk queries to the specific H3Africa projects, but to other members of the African research community. The SAB also recommended seconding an individual from each Node to be part of the helpdesk and thereby increase the profile of the helpdesk within those Institutions.

As with all other working groups, attendance at the USWG meetings and follow up on assignments is still irregular with certain core groups shouldering numerous tasks that results in a slow turnaround time for some of the deliverables. The following were suggested for consideration:

- Get more people recruited to the USWG from partner institutions, that way there is institutional/node representation at each session.

- Consider dropping habitually absconding members.

- Have a rotating rapporteur to follow up issues for each session.

- Cut back on meet frequencies to once a month

Another goal suggested that the USWG should explore and facilitate is to help provide access to publically available data and determine what types of data the H3Africa consortium and H3ABioNet members would like.

Dr. Judit Kumuthini.   Dr. Jonathan Kayondo.

H3ABioNet Montly Newsletter Issue 6 : October 2013

Node Accreditation

- **Node Accreditation**

During its first year of activity, the NAWG was able to prepare assessment exercises for the two types of datasets most likely to be produced by the H3Africa research projects: high-throughput sequencing of human exomes or full genomes, and genotyping of large cohorts of patients and unaffected controls followed by GWAS analysis. For both types of analysis, the NAWG has produced synthetic datasets or subsets of public datasets that will guarantee that all candidate nodes will tackle distinct problems, but that they will be of essentially identical complexity. Standard Operating Procedures were written up to provide guidance on current standard for the analysis of the data, and an overall Protocol for the administration of the assessments explains the way that the tests will be administered and evaluated. The NAWG is now ready to assess candidate nodes.

At the suggestion of the H3ABioNet Scientific Advisory Board, and in discussion with representatives of many network nodes, the focus of the NAWG will shift during its second year of operation to assessing skills that are not necessarily linked to the needs of the H3Africa research projects but are deemed important to documenting the general bioinformatics capabilities of the nodes. The exact list of these new tests has not been finalized, but it will most likely include:

1. Analysis of 16S rDNA amplicons from bacterial communities and derivation of the OTUs present in the community
2. Assembly and annotation of relatively simple (up to 200 Mbases) genomes
3. Analysis of microarray-based gene expression profiling experiments, including derivation of differentially expressed gene lists
4. Processing and analysis of RNA-seq data,
5. Production of multiple sequence alignments and derivation of phylogenetic relationships.

This extension of the role of the NAWG will allow a much wider proportion of nodes to be assessed for their competence in bioinformatics' data analysis.

The reservation with trying to implement a wide variety and number of SOPs from the NAWG's perspective is the preparation of the assessment exercise together with the SOPs and synthetic datasets is a non-trivial exercise for which mainly the University of Illinois and Central has been providing the main contributions. Hence the suggested SOPs will be ranked according to the priority and expertise the groups have at present to construct such SOPs and datasets.

So far a couple of Nodes have expressed interest in taking part in the accreditation exercise for the GWAS SOP that has been developed. Additionally Nodes that have taken the assessment will have the details entered within NetCapDB to indicate their ability to perform a specific analysis. The Node accreditation exercises will also be included for the 5 year developmental plan a Node will have to submit.

Dr. Victor Jongeneel.

- **H3ABioNet Second Annual meeting**

H3ABioNet Second Annual Meeting – Johannesburg 2013

The second annual H3ABioNet meeting was held in Johannesburg, South Africa from the 1st to the 3rd of October to coincide with the third H3Africa consortium meeting. The second annual H3ABioNet meeting was attended by 41 people including 2 scientific advisory board members. Day 1 of the meeting was reserved for discussion and progress updates of all the H3ABioNet Nodes. Some notable successes were achieved such as the opening of the Egyptian Center for Bioinformatics and Genomics (ECBAG), which is a first for Egypt, to various bioinformatics training workshops and equipment of a computer training lab by the Ghanaian and Kenyan Nodes. A challenge that all the Nodes have been facing is the recruitment of suitably qualified personnel to fill bioinformatics vacancies available at their institutions although they have been some success stories of African scientists trained in Europe returning to their home countries due to the opportunities offered by H3ABioNet with Niger Node having managed to recruit the country's first PhD in bioinformatics. The main activities that the Central Node has been involved in were reported to the GA by Prof. Nicola Mulder.

The various working group Chairs provided feedback to the GA on their various progresses and challenges faced with the mode of meeting communication being a major concern. Central did spend the beginning of the year looking at various communication platforms with Adobe connect being the preferred choice and is currently being used by the ISWG. However, due to some initial issues during the attempted implementation of the Adobe connect platform for other working group meetings and its resulting unpopularity, Skype has remained as the default meeting platform. The Institutional communication infrastructure at Central is currently being migrated to VOIP with the widespread adoption of Microsoft Lync being rolled out. Currently, UCT will only shift to Lync 2013 which has clients for a variety of platforms (Unix / Mac based) in 2014. Central will test this application to determine if it does provide better call quality and subsequently adopt Lync for future working group meetings if successful.

Another challenge faced by the working groups is the non-attendance of members which hampers any activities a H3ABioNet working group is trying to implement as all the responsibilities for completing certain tasks are borne by limited number of individuals that consistently attend the meetings. In some cases, membership of a working group consist of PIs who have very limited time and can make minimal contributions. It was decided that the PIs should get their H3ABioNet funded students, postdoctoral and other researchers involved with the working groups to help in their skills development and make meaningful contributions while reporting to the Node PI's on the current progress of the working group they are part of.

A brief overview of the NetCapDB system that is being implemented to assess the impact of the NIH capacity building exercise and facilitate annual reporting was presented to the GA. The main issue for the Nodes was the amount of baseline data

- **H3ABioNet Second Annual meeting**

H3ABioNet Second Annual Meeting – Johannesburg 2013

to enter within NetCapDB which spans 5 years before the project was initiated. However, if bioinformatics capacity within Africa is to be measured before the start of H3ABioNet, a good baseline indicator that spans the project's duration is required as well as to smooth out any anomalies e.g in a certain year a Node might graduate 10 PhD students or produce 15 publications while in a different year, the same Node might produce no publications or have any students graduating. The first scenario will indicate that the capacity existed and has not been improved upon by the project and the second scenario might indicate that the project is roaring success, neither which would be a true reflection on H3ABioNet's impact due to a sampling bias. Some members did find the NetCapDB interface for entering data to be well structured and easy to use and requested if it can be adopted for other projects. A valuable point raised by Dr. Julie Makani is that for NetCapDB to measure H3ABioNet's impact, other non H3ABioNet bioinformatics institutes and labs should enter their own data in to provide a control when measuring the impact of H3ABioNet.

The Nodes are required to provide annual reports and complete a 5 year plan for developing their bioinformatics capacity for which a suitable outline was presented and discussed by the GA. Various points were suggested by the PIs which will be adopted into the template to be completed by the Nodes.

Each working groups' Chairs presented a review of the year 1 milestones and the year 2 milestones which were discussed in depth with some being reassigned and new dates incorporated to the year 2 milestones from the specific projects. The day ended with each PI of a specific H3Africa working group providing feedback from that H3Africa working group. It was determined that attendance to the H3Africa meetings has been patchy which is problematic for H3ABioNet as there is no flow of information from the various H3Africa working groups to H3ABioNet, making it impossible to plan for data acquisitions and storage. Hence, an audit of the H3Africa working groups participation and the H3ABioNet working groups will be done with the various Nodes being asked to assign an active participant to each working group and those that do not attend / participate will be asked to step aside to enable other members to participate.

Nicki Tiffin and Jean-Baka hard at work on day 1 of the second H3ABioNet annual meeting

- **H3ABioNet Second Annual meeting**

H3ABioNet Second Annual Meeting – Johannesburg 2013

Day 2 of the second annual H3ABioNet meeting invovled a review of the H3ABioNet project by the scientific advisory body. Unfortunately two of the SAB members were unable to attend the meeting due to visa issues and one being an NIH employee and all NIH staff were recalled back to the USA due to their federal government shut down. Professor David Roos and Professor Jessica Kissinger were present to provide invaluable feedback based on their extensive experience working in large consortiums and networks. Professor Nicola Mulder presented a 30 minutes overview of the H3ABioNet consortium, its activities and future plans for year 2. The working group Chairs presented their progress to the SAB and received feedback as reported in the individual working group sections of this newsletter. Specific research projects were also presented to the SAB who provided feedback on the various presentations as summarised in the RTDWG section.

Professors Jessica Kissinger and David Roos from the Scientific Advisory Body grilling the H3ABioNet PIs and providing invaluable feedback.

A series of specific questions we poised to the SAB which included "How to get the H3Africa PIs to engage with H3ABioNet?". The SAB suggested that Ambassadors from the various Nodes should be appointed that will liase directly with the H3Africa PIs in order to engage them to determine their needs and chase up on outstanding documentation required and will form the first point of contact to followup on. Other operational details pertaining to finance and communication were discussed at length as well (please refer to the meeting minutes available here).

The third day of the meeting overlapped with the third H3Africa consortium meeting for which all the H3ABioNet working groups had their meetings and developed strategies for upcoming year.

H3ABioNet Montly Newsletter Issue 6 : October 2013

- **H3ABioNet Second Annual meeting**

H3ABioNet Second Annual Meeting – Johannesburg 2013

Day 2 of the second annual H3ABioNet meeting with the PIs and SAB

Group photo of H3ABioNet second annual meeting particpants and SAB

Slyvester, Pandam and Winston taking a break from interacting with the H3Africa consortium PIs and catching up on some email

H3ABioNet Second Annual Meeting – Johannesburg 2013

- **H3ABioNet Second Annual meeting**

H3ABioNet members Ayton, Gerrit and Victor discussing SOPs and other bioinformatics related topics

Dr Julie Makani and Prof. David Roos exchanging notes during the coffee break on day 3 of the H3ABioNet second annual meeting

Prof. Nicola Mulder presenting the H3ABioNet helpdesk and progress on the Node Assessment exercise to the H3Africa PIs during the 3rd H3Africa consortium meeting

- **Announcements**

Announcements

- **Congratulations to Dr. Özlem Tastan Bishop and the Rhodes University Bioinformatics Research Group has been officially recognised as an official Unit by Rhodes University – well done to Dr. Özlem and the Rhodes University Bioinformatics Research Unit!**

- **The Egyptian Node published a conference paper at the Advances in Computing, Communications and Informatics (ICACCI), 2013 titled: "Using formal concept analysis for mining hypomethylated genes among breast cancer tumors subtypes" which is available here.**

- **A work shop has been organized in Zagazig University, Egypt by the Egyptian Node, titled "The Utilization of Bioinformatics and Genomics analysis methods in biological experiments", as one of the activities of The Egyptian Center of Bioinformatics and Genomics (ECBAG) at the faculty of Agriculture. The workshop took place on 27/10/2013 and many young scientists and scholars have attended the work shop. The workshop included some lectures about H3BioNet and many lectures about current bioinformatics research in the center and possible cooperation between different institutions in Egypt and Africa. The researchers in the Egyptian node have presented lectures on different topic in Bioinformatics such as "Introduction to bioinformatics", "Biomathematics", " Medical Image Analysis", "Biological Sequence Analysis", "Modeling" and "Next Generation Sequencing". This workshop was funded by the faculty of Agriculture, Zagazig University and a private company   Analysis". The ECBAG team is planning to organize many more workshops and training courses on Bioinformatics in the future.**
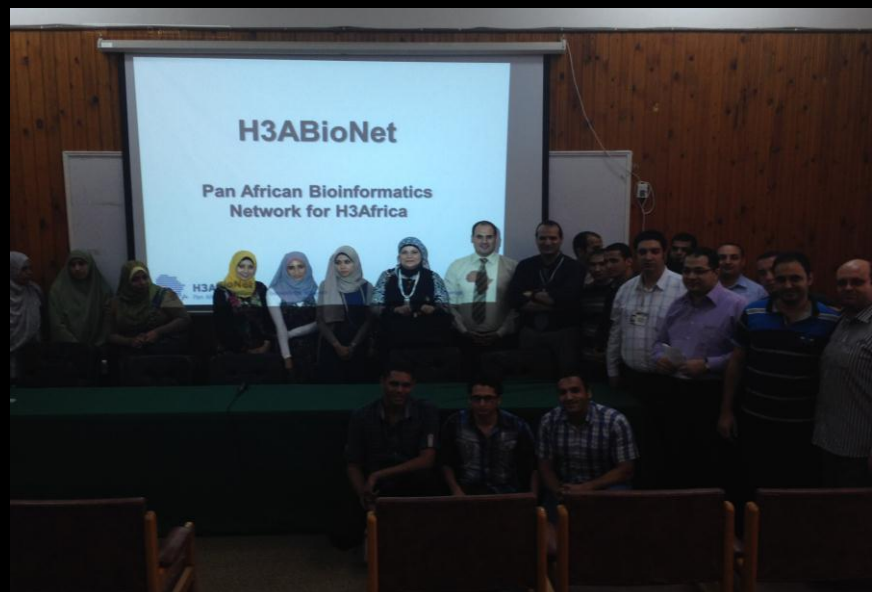
Zagazig University Bioinformatics Workshop participants trained by the Egyptian Node

H3ABioNet Montly Newsletter Issue 6 : October 2013

- **Announcements**

Important Dates

- 31st October: All Nodes to submit their financial reports to Central

- 14th – 15th November: Hands on NGS workshop hosted by CPGR at the Centre for High Performance Computing, South Africa

- 25th November: Submission of progress report for the 18month report to the NIH

- 26th November: Bioinformatics Support Platform Stakeholders meeting South Africa

- 30th November: 18 month report to NIH