Acknowledgments

The node managers would like to thank the following people for their support in the course:

- Harald Dahle, for setting up the student accounts and the environment on the EMBnet Norway side.
- Alfredo Hernández Alvarez and the technicians of the LCG UNAM IT unit for recording the course and making the laboratory equipment run smoothly for the entire duration of the course.

References

1. Hekkelman M L and Vriend G. (2005) MRS: a fast and compact retrieval system for biological data. Nucleic Acids Research 33 (Web Server issue):W766-W769
2. Hull D, et al. (2006) Taverna: a tool for building and running workflows of services. Nucleic Acids Research 34 (Web Server issue):W729-W732
3. Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA. URL: http://accelrys.com/products/pipeline-pilot/
4. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics 16: 276-277

EMBRACE Workshop - "Next Generation Sequencing II"

Rome, Italy, November 2009

Andreas Gisel[1], Erik Bongcam-Rudloff[2]

[1]Institute for Biomedical Technologies, [2]Hgen, SLU/UU, SE

After the success of the EMBRACE [1] workshop on "Building Next Generation Sequencing platforms and pipeline solutions", in Rome in November 2009, with more that 60 international participants [2], the organization committee (Domenica D'Elia, Erik Bongcam-Rudloff, Andreas Gisel) decided to organize a second event to keep pace with the fast development of Next Generation Sequencing (NGS) technologies, and the tools and methods necessary to analyse and interpret NGS data.

With financial support from the NoE EMBRACE, the organization committee decided to organize the second NGS workshop in conjunction with the EMBnet Annual General Meeting (AGM) in Ruvo di Puglia in June 2010 [3].

While the workshop in Rome focused on the basic problems arising from the use of NGS technologies, such as data management and storage or mapping and assembly of NGS data, the workshop in Ruvo di Puglia had a different focus,

Figure 1. Workshop participants during the Hackaton

Figure 2. Workshop participants

highlighting areas where NGS technology is useful, such as epigenetics, small RNA analysis and reference guided assembly strategies.

The organisation committee was again able to invite talks from high-profile scientists from all over Europe, to give this event the importance it deserves.

Peter Rice, from the European Bioinformatics Institute (EBI) in Hinxton, presented developments of the EMBOSS [4] software suite to facilitate NGS data analysis. EMBOSS will cope with the new data formats and the tremendous data volumes.

Stefan Marklund, from the Swedish University of Agricultural Sciences in Uppsala (SLU) [5], demonstrated the potential of NGS in epigenetics research, accelerating investigations on a genome wide scale and our understanding of methylation events and their consequences.

Bastien Chevreux, from DSM Nutritional Products AG in Switzerland, presented the powerful assembly suite MIRA [6] and some strategies on how to successfully assembly denovo sequenced bacterial genomes.

Alberto Policriti, from the University of Udine and Applied Genomics [7], Italy who spoke at the previous workshop in Rome, this time presented a very interesting sequence-assembly approach, combining denovo assembly and mapping onto a similar reference as guidance.

Rasko Leikonen, from the European Nucleotide Archive at EBI in Hinxton, presented developments of the Sequence Read Archive (SRA, 8), the necessity for researchers to continue to submit their NGS data as they normally do with single sequences to EMBL, NCBI or DDBJ. The SRA was presented to the NGS society by Guy Cochrane during the Rome workshop.

David Horner, from the University of Milan, Italy, introduced the participants to the world of small RNAs, and demonstrating the potential of NGS technology to understand this new world and its effect on different organisms.

Finally, Massimiliano Gentile from the IT Center for Science (CSC) in Helsinki, Finland, explained the use of NGS technologies for chromatin-immuno-precipitation combined with sequencing, and its power to search for new transcription

factor binding sites. Further, he demonstrated CHIPSTER [9], a user-friendly analysis and workflow tool, originally developed for microarray data analysis, now upgraded to handle NGS data. The ChIP-seq data analysis was the first example to show the potential CHIPSTER has.

The presentations took place during the first day of the two-day workshop, and were followed by a sight-seeing tour to the cathedral of Trani, to the famous castle of Frederic II in Castel del Monte, and finally to a well-earned dinner at the Masseria San Giovanni at Altamura. The whole social event was generously sponsored by the Apulia Region [10]].

The second day of the workshop was dedicated to a sort of hack-a-thon, where participants were able to put their hands on algorithms and real NGS data, in order both to understand the various problems in NGS data analysis and to stimulate discussions and collaborations in various fields of NGS technology.

Rasko Leinonen and Vadim Zalunin prepared a hands-on session on how to submit and retrieve NGS data from the SRA, Bastien Chevreux guided the participants through the functionalities of MIRA3, and Alberto Policriti and Francesco Vezzi prepared a session with the read assembler, Velvet [11], the read mapper SOAPdenovo [12], and the Enhanced Reference Guided Assembly (e-RGA) pipeline.

At the end of the event, Erik Bongcam-Rudloff presented to the participants an EU COST-Action [13] proposal that he was coordinating, to build an NGS data-analysis network. The proposal passed the first stage and Erik invited the participants to join preparations for the full proposal.

The event hosted 68 participants from 30 countries all over the world, demonstrating that the topic of NGS data analysis is very important in the bioinformatics and bioscience communities.

Acknowledgements

We would like to thank the institutions who made this workshop possible through their support: in particular, the European Commission through its funding of the EMBRACE Network of Excellence (the EMBRACE project is funded by the European Commission within its FP6 Programme, under the thematic area "Life sciences, genomics and biotechnology for health", contract number LHSG-CT-2004-512092); the Regional Council of Apulia (IT)(initiative financed by EU funds POR Puglia 2000/2006); the CNR Institute for Biomedical Technologies of Bari (IT); the National Institute of Nuclear Physics of Bari (IT); and EMBnet (the European Molecular Biology Network).

The organizers would also like to express their deep gratitude to all the invited speakers and all the participants who, through their enthusiasm and interest, made this such a successful workshop.

Figure 3. Workshop participants during the Hack-a-thon

References

1. http://www.embracegrid.info
2. www.nextgenerationsequencing.org
3. http://www.embnet.org/NGS-AGM2010
4. www.emboss.org
5. www.slu.se
6. http://sourceforge.net/apps/mediawiki/mira-assembler
7. https://www.appliedgenomics.org/
8. http://www.ebi.ac.uk/ena/
9. http://chipster.sourceforge.net/
10. http://www.viaggiareinpuglia.it/aptbari
11. http://www.ebi.ac.uk/~zerbino/velvet/
12. http://soap.genomics.org.cn/soapdenovo.html
13. http://www.cost.esf.org/