

Editorial

Special Issue: Next Generation DNA Sequencing

Paul Richardson

PRC LLC, 25 Amber Lane, Lafayette, CA 94549, USA; E-Mail: paulrichardson07@gmail.com;
Tel.: +1 925 301-5460

Received: 21 October 2010 / Accepted: 26 October 2010 / Published: 27 October 2010

Next Generation Sequencing (NGS) refers to technologies that do not rely on traditional dideoxy-nucleotide (Sanger) sequencing where labeled DNA fragments are physically resolved by electrophoresis. These new technologies rely on different strategies, but essentially all of them make use of real-time data collection of a base level incorporation event across a massive number of reactions (on the order of millions *versus* 96 for capillary electrophoresis for instance). The major commercial NGS platforms available to researchers are the 454 Genome Sequencer (Roche), Illumina (formerly Solexa) Genome analyzer, the SOLiD system (Applied Biosystems/Life Technologies) and the Heliscope (Helicos Corporation). The techniques and different strategies utilized by these platforms are reviewed in a number of the papers in this special issue. These technologies are enabling new applications that take advantage of the massive data produced by this next generation of sequencing instruments.

In this special issue, nine papers review and demonstrate the utility and potential of next generation sequencing. One of the biggest consequences with NGS technologies is how to deal with all of the data produced by these platforms. Magi *et al.* [1] review the software tools available for the multiple functions needed to process and interpret the huge amounts of data produced by these instruments. Sequence alignment to a reference, polymorphism detection, *de novo* assembly and visualization software are covered in this paper. Knudsen and colleagues [2] introduce a computer simulator that utilizes real and simulated reads to assess the effects of different factors and strategies for utilizing NGS to perform *de novo* assemblies. This is presently a particular challenge for the generally short (sometimes unpaired) reads produced by NGS Technologies. Another review article [3] on bioinformatic issues with NGS data focuses on statistical methods for analysis of Chip-Seq data as well as RNA data generated by these methods. NGS platforms are particularly suited to the use of ChIP-Seq methods as an alternative to ChIP-ChIP methods for identification of transcription factor binding sites.

Another application that is particularly well-suited for the types of data produced using NGS is microRNA sequencing. MicroRNAs (miRNAs) and short interfering RNAs (siRNAs) are small RNA molecules of 17 to 24 bases that have been shown to play a critical role in gene regulation by mediating RNA interference. Motameny *et al.* [4] discuss in a review article the methods and strategies for capturing and sequencing miRNAs and their analysis. Epigenetic modifications, especially in the form of DNA methylation is another area of gene regulation that has recently gained significance in the research community. NGS platforms offer unprecedented abilities to elucidate the state of methylation in different species, individuals, tissues and cell types. Recent advances and strategies for characterizing the methylome are covered in two review articles [5,6] in this special issue.

The subjects of metagenomics (environmental sequencing) and sequencing of ancient DNA samples are two additional areas where NGS platforms have provided unforeseen opportunities for generating previously unattainable levels of understanding from these samples. Rooks *et al.* [7] demonstrate the utility of the 454 platform for assessing the viral ecology of a specific freshwater environment. Edwards *et al.* [8] utilized 454 technology on a marine microbial biofilm community to identify species and genes involved in cellulose degradation. Knapp and Hofreiter [9] review the progress and strategies for utilizing NGS platforms for generating genome sequence from ancient DNA samples, which are often contaminated with non-target DNA. The relatively inexpensive production of large amounts of data allows researchers to zero in on those sequences of most interest. Recent examples covered include woolly mammoth, Neanderthal and Palaeo-Eskimo. Finally, Anderson and Schrijver [10] review different NGS platforms and discuss the impact on the future of genomic medicine.

The articles in this special issue represent a cross section of applications demonstrating the utility of NGS. New sequencing platforms are being introduced (next next gen, or third generation sequencing) by companies utilizing single molecule detection techniques to record individual DNA molecules during incorporation (Pacific Biosciences) or reading DNA directly by sieving it through nanopores. It is an exciting time in the field of genomics with rapidly changing and advancing technologies. The costs of sequence will surely continue to decrease and new applications of this technology will be introduced to address critical areas in academic and clinical research. I would like to thank all the authors and reviewers who contributed to this special issue and look forward to hearing about new sequencing technologies and application advances in the future.

References and Notes

1. Magi, A.; Benelli, M.; Gozzini, A.; Girolami, F.; Torricelli, F.; Brandi, M.L. Bioinformatics for Next Generation Sequencing Data. *Genes* **2010**, *1*, 294-307.
2. Knudsen, B.; Forsberg, R.; Miyamoto, M.M. A Computer Simulator for Assessing Different Challenges and Strategies of *de Novo* Sequence Assembly. *Genes* **2010**, *1*, 263-282.
3. Ghosh, D.; Qin, Z.S. Statistical Issues in the Analysis of CHIP-Seq and RNA-Seq Data. *Genes* **2010**, *1*, 317-334.
4. Motameny, S.; Wolters, S.; Nürnberg, P.; Schumacher, B. Next Generation Sequencing of miRNAs – Strategies, Resources and Methods. *Genes* **2010**, *1*, 70-84.
5. Zhang, Y.; Jeltsch, A. The Application of Next Generation Sequencing in DNA Methylation Analysis. *Genes* **2010**, *1*, 85-101.

6. Taylor, K.H.; Shi, H.; Caldwell, C.W. Next Generation Sequencing: Advances in Characterizing the Methylome. *Genes* **2010**, *1*, 143-165.
7. Rooks, D.J.; Smith, D.L.; McDonald, J.E.; Woodward, M.J.; McCarthy, A.J.; Allison, H.E. 454-Pyrosequencing: A Molecular Battiscope for Freshwater Viral Ecology. *Genes* **2010**, *1*, 210-226.
8. Edwards, J.L.; Smith, D.L.; Connolly, J.; McDonald, J.E.; Cox, M.J.; Joint, I.; Edwards, C.; McCarthy, A.J. Identification of Carbohydrate Metabolism Genes in the Metagenome of a Marine Biofilm Community Shown to Be Dominated by Gammaproteobacteria and Bacteroidetes. *Genes* **2010**, *1*, 371-384.
9. Knapp, M.; Hofreiter, M. Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. *Genes* **2010**, *1*, 227-243.
10. Anderson, M.W.; Schrijver, I. Next Generation DNA Sequencing and the Future of Genomic Medicine. *Genes* **2010**, *1*, 38-69.

© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).