MAR 550-4, Fact sheet 1

Deep Sequencing

Definition

There is no readily available definition of deep sequencing in the literature, but it is apparent that deep sequencing is unique from shallow sequencing. The confusing is most likely from recent advancements in next generation sequencing (NGS) (Czerniecki and Wołczyński 2011). Using Sanger sequencing techniques, sequencing a genomic region ~10 times may constitute as deep sequencing. With NGS, deep sequencing may mean sequencing a genomic region ~100s-1000s of times. The overall difference between shallow and deep sequencing is coverage (Illumnia, 2015b). Deep sequencing provides incredibly extensive coverage for a target region of the genome, allowing for identification of rare clonal types, transcript isoforms, single nucleotide polymorphisms (SNPs), and novel exons.

Methods

Deep sequencing often generates short reads, and can generate millions of short reads if using NGS methods. Then, these short reads are mapped to a reference genome; a similar procedure to full genome assembly, but only for a portion of the genome. Conventional tools like BLAST and BLAT do not work for deep sequencing data, but there are other tools available to process this vast amount of data to align short reads, assemble overlapping reads, then identify SNPs within the target genomic region (Benton et al., 2009). This analysis can be carried out through an interactive software or through a uniquely programmed pipeline.

Applications

Ribosome profiling is a technique based on deep sequencing ribosome-protected mRNA fragments within a genome. This can give a snapshot of sequences being actively transcribed within a eukaryotic cell at a specific time (Illumina, 2015a). Ribosome profiling can also be used to observe *in vivo* translation and identify novel exons that arise from transcript isoforms (Ingolia, 2014). By using ribosome profiling and deep sequencing, one can get closer to identifying the translated proteins from a genomic sequence. Deep sequencing can also be used to identify rare clonal cell types that compromise as little as 1% of the sample. This feature is crucial in identifying mutations within tumor because tumors can contain normal cells in addition to multiple sub-clones of cancer cells with mutated genomes (Illumina, 2015b). Accurate identification of these mutations can guide cancer research and unique genetic-based medical treatments.

References

Benton, A., Darko, S., Prakriti, M., Takundwa, P. (2009). "Introduction to Bioinformatics Seminar." http://docslide.us/documents/sam5472de47b4af9f844c8b45c2.html

Czerniecki, J., & Wołczyński, S. (2011). Deep sequencing–a new method and new requirements of gene expression analysis. *Studies in Logic, Grammar and Rhetoric*, *25*(38), 41-47.

Illumina (2015). "An Introduction to Next-Generation Sequencing Technology." http://www.illumina.com/technology/next-generation-sequencing.html

Illumnia (2015). "Next-Generation Sequencing: Deep Sequencing." http://www.illumina.com/technology/next-generation-sequencing/deep-sequencing.html

Ingolia, N. T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nature Reviews Genetics*, *15*(3), 205-213.