Busra Nur Ata
28749
Spring 2020

Tsai et al. 2013

Summary: This article gives an overall idea about the NGS technology and introduces the tools that are used to process the raw data. NGS technology is cheap fast and accurate technology which is used to sequence mainly whole genome DNA, RNA sequencing and small RNA sequencing. Firstly, any NGS raw data image is converted to FASTQ format files. Then reads are obtained after clusters are formed and low-quality reads are eliminated. Read alignments are done according to reference sequences. De novo alignments are used to identify a new genome or RNA sequences. From short reads, significant contigs are identified. Single Nucleotide Variations (SNVs) can also be detected by NGS. Additionally, large DNA alterations such as indels, copy number variants, inversions and translocations can be detected by NGS DNA Sequence Analysis. RNA-seq is preferred dominantly to profile gene expression levels. Counting reads per transcript provides a differentially expressed gene profiles. Transcriptomic analyses also contribute to epigenetics by identifying the RNAs that is involved transcriptional regulations. Finally, small RNAs analysis can be performed after cDNA library reads is obtained and adapter part in the reads are eliminated. sRNA-seq analysis helps to identify miRNAs which are reported as important due to its possibility of being a biomarker for diseases and its potential for the drug development.

Question: Why expression distribution of RNA-seq data is different from conventional microarrays?

RNA-seq measures all RNA transcript counts, a direct measurement of gene expression, while the microarray measures a fluorescence intensity that is due to hybridization with anti-sense probe sequences, an indirect measurement of gene expression. The advantage of RNA-Seq over microarrays is that it provides an unbiased insight into all transcripts.

Ong et al. 2019

Summary: Article highlights the whole exon sequencing (WES) and its impact on personalized medicine. WES detects the variations in the exon regions and finds causal variants of a disease or disease-causing mutations. WES is cheaper than whole genome sequencing (WGS) since only exons are read and analyzed and coverage is higher, which provides to detection of rare variants. WES is examined under 10Vs. Briefly, volume of the data for WES is less than WGS with same coverage. Generation of the reads per sample is faster for WES data. Those features of WES make it more appealing towards WGS. However, it is guessed that cost reduction in WGS may lead to shift to perform WGS analysis rather than WES since every disease-causing variant doesn't occur on the exons. The article also gives a brief information about tools to analyze NGS data. At the end, application of the artificial intelligence on the personalized medicine is mentioned and technology companies are interested in developing tools that helps to understand the data by analyzing from different aspects.

Question: Why did GWAS studies not give satisfactory relation between diseases and the variants?

Although GWAS studies are comprehensive, disease related variant is not the only reason for the disease. For most cases, GWAS found the risk factors and metabolic pathways related to multifactorial diseases.

Busra Nur Ata
28749
Spring 2020

Salk and Kennedy 2019

Summary: This article focuses on the detecting mutagenesis by using NGS. Since mutagenesis is one of the driving forces for carcinogenesis, detection and evaluation of the mutagenesis on the DNA is important. There are in vitro and in vivo assays to detect the mutations, which provide limited mutation detection. NGS enhances the detection of low-frequency somatic mutations and DNA damage. However, NGS has a high error rate that leads a confusion about presence of mutation on the DNA fragment. Bioinformatic filtering and enzymatic removal of DNA damage are some improvements to decrease false variant calls. Also, error-connected NGS is developed to increase accuracy, which is based on consensus-based error connection.

Somatic mutations may give a pattern to understand whether the mutation triggers cancer or not. Background error rate of sequencing helps to understand the status of the cells. Therefore, NGS platform is used to strengthen the relation between mutagenesis and carcinogenesis.

Question: The scientist has received an NGS reads from somatic single cell types. When s/he aligned the reads, s/he has realized that almost half of the reads have the same somatic mutation at specific point compared to reference sequence. What do you comment about this mutation?

Since half of the reads has somatic mutation, genotype of the person is a heterozygous, which means it is an allele.