

MULTIOMICS ANALYSIS

Overview

The repository of data available in Bioinformatics spans from genomic, Transcriptomic and proteomics according to the central dogma of biology. Where different representation can give multiple perspective on the functional and structure of a spans of data available .To explore complex biological process definite manner ,It is vital for an integrated approach for understanding the interdependence of the molecules and function .

Bioinformatics is standing paradigm of transferring a study to a large scale research,Also With the mammoth data available in the bioinformatics we as a company give more insight with the kind of analysis and the workflow and method followed for Multi Omics analysis in a efficient and streamlined manner with the resources and computational capacity

Challenges & SAMPLE DATA

Challenges :

With the mammoth of data available for Multi omics it is time and computationally expensive

Sample data :

The workflow that is being followed for Multi omics analysis is explained for 150 WGS samples that was received from one of our clients for analysis,Using the data the downstream analysis for the span of Multi omics data

APPROACH

- In the Multi omics analysis the downstream analysis the data considered is Genome and transcriptome and respective plots are created .The beginning point of the WGS analysis is the raw FASTQ files provided
- For Genomic analysis the variants information is obtained from the Variant annotated file and the output is given category of SNV,CV,Indels and CNV
- For Transcriptomic analysis the gene expression is obtained and the TPM and FPKM values are added as a visual representation of heatmap

Genomic analysis

GENOMIC ANALYSIS :

- Conversion of bam from FASTQ file provided
- Creating VCF files from bam files provided
- Annotating the files with VeP after the set of filters based on Variant quality, Mapping, Depth
- Variant analysis is carried out ,Output data that is falls in the category of
 - SNV
 - CV
 - CNV
 - Indels
- The data obtained from such variants has been filtered for the disease of interest and then the entire variants is filtered for the same under different categories mentioned above
- Only those variant falling in those gene of interest is taken and filtered for the same
- The variants that falls under the cancer of interest and gene of interest is visualised using different plots are attached for WGS data

WORKFLOW

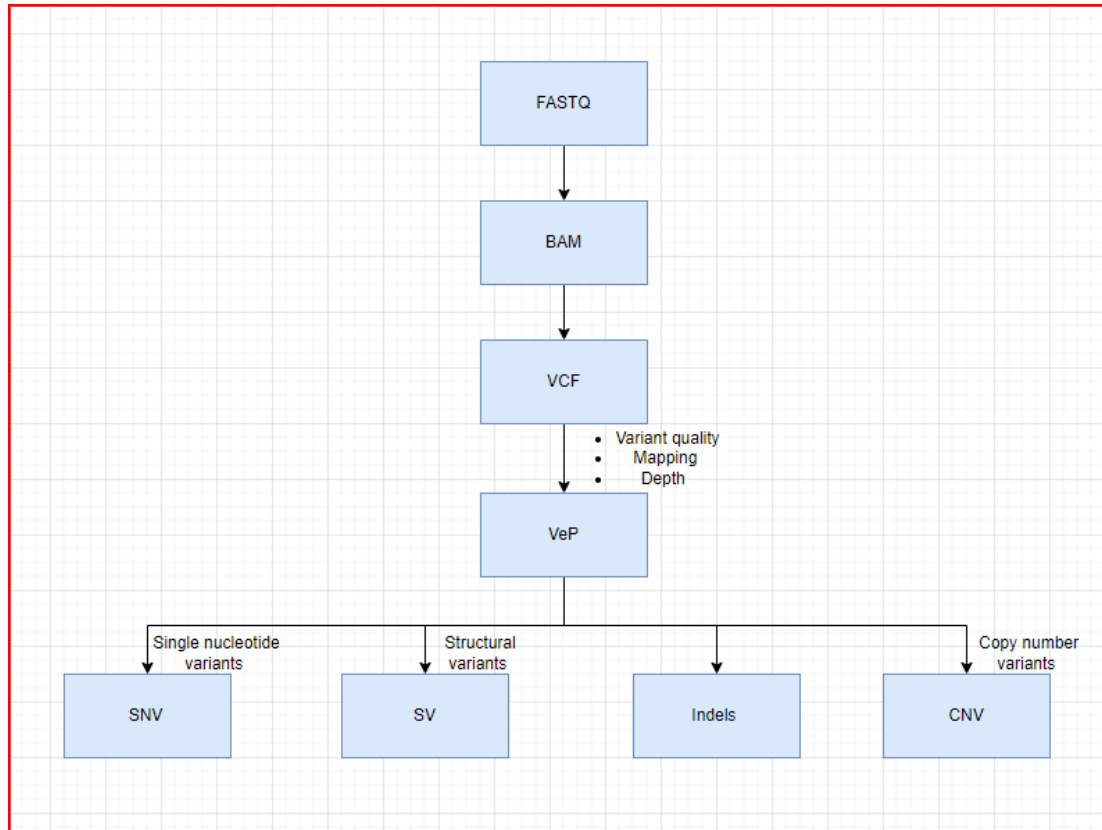


Fig 1: Workflow for the genome analysis

RESULTS

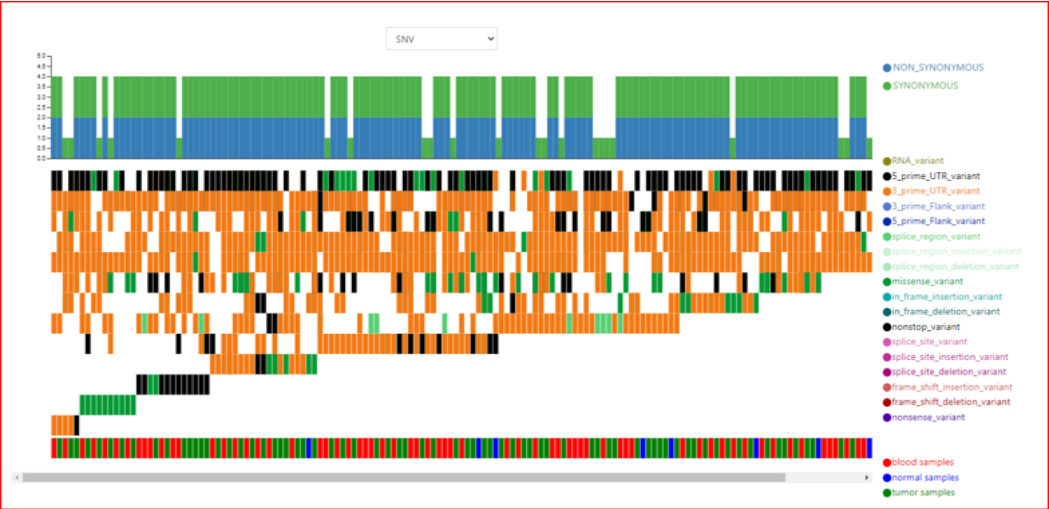


Fig 2: Consequence of the variants available from WGS data

RESULTS

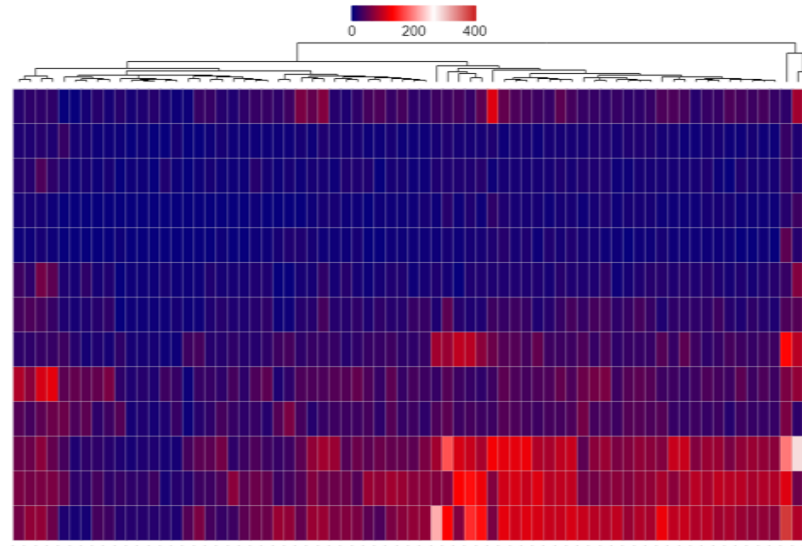
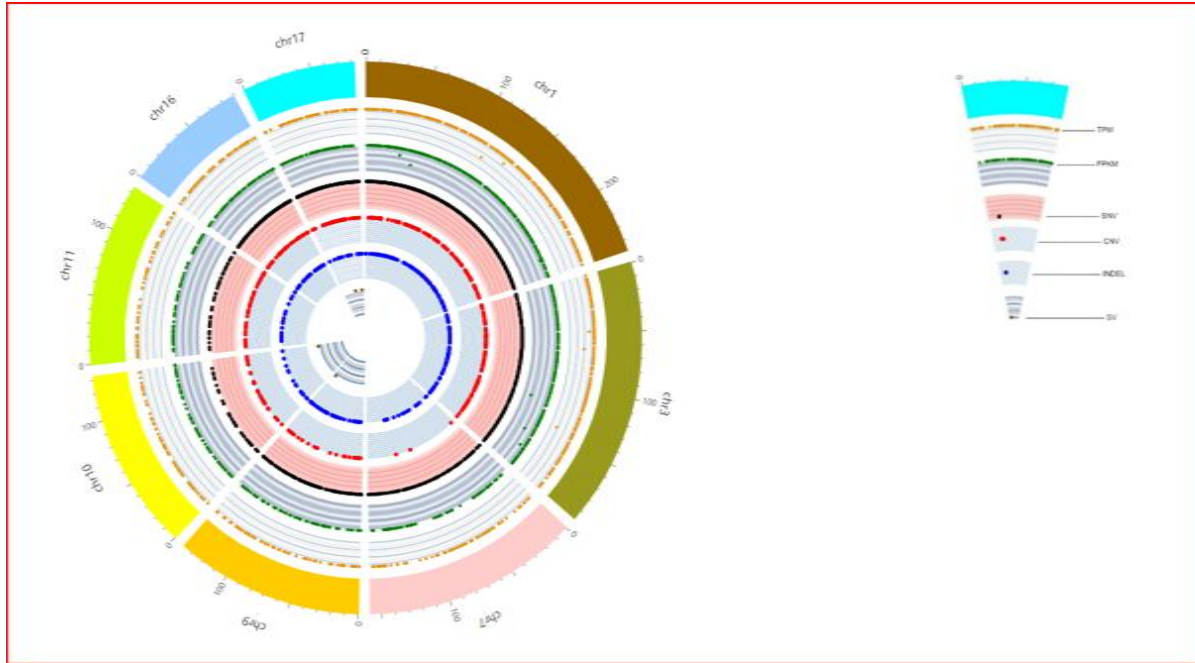


Fig 3 : Heatmap for the FPM and TPKM values from the gene expression

MULTI-OMICS ANALYSIS PLOT



The data from the plot we can see that the chromosome level information is given in hierarchy of Transcriptome and genome level data in a Circos plot